

(An) Impact Roadmap

Overview

Executive Summary

1. [Key Decisions \(and Tools\)](#) - The most important decisions we'll make, and how to make them

There are some useful [tools](#) that it makes sense to learn early on, as they can be used throughout the roadmap. Not least, to help you with some of the most important decisions you'll make on the path to impact, such as defining your [values](#), and choosing your [career area](#) based on personal fit. If starting a nonprofit is the best fit for you, then we can move on to deciding on a [charity idea](#), finding a [co-founder](#) and selecting your (initial) implementation [country](#). We can now begin to more concretely understand our path to impact through an (initial) Theory of Change.

2. [Theory of Change](#) - How and why the program will work

Now that you've made your key decisions, you can move on to defining your potential program. You'll do this by first outlining some [Project Objectives](#), before creating your first (likely somewhat broad and vague) [Theory of Change diagram](#), sketching out the Activities you will undertake which lead to your projects Outputs, resulting in the overall positive Outcomes of your project, which ultimately achieves Impact. This will provide some clarity around the impact you intend to make, as well as the pathway to that impact. It's likely though that there are some gaps in your knowledge and some uncertainties and assumptions in your Theory of Change, which brings us to an Evidence Review.

3. [Evidence Review](#) - A critical review of secondary sources related to our program

We now have an idea of how we intend to make an impact. It's time to review any evidence that could help us more concretely answer some of our uncertainties and clarify our assumptions. To do this, we'll undertake a [literature review](#) to help us answer these questions (i.e. What interventions have been tested or studied in relation to this program? Under what conditions were they implemented? Were they effective? Why or why not?). We may even start [contacting experts](#) to further calibrate our understanding. Still, however, we will need to get a better understanding of our chosen context, so we'll go on to undertake a Needs Assessment.

4. [Needs Assessment](#) - Seeking out perspectives from the populations we aim to help

We now have a good overview of the relevant evidence in our field, but it's time to get a deeper understanding of [the context](#) we aim to work in. This will involve gathering inputs and perspectives from the community, and more specifically, helping us (again) answer any uncertainties and clarify any assumptions within our Theory of Change. This will involve detailing what we want to learn, and how to gather data to support this. After this, there may still be questions you don't know the answers to, and may need to engage in [Primary Research](#) of your own. Either way, after your Needs Assessment is complete, you'll be moving on to developing your intervention.

5. [Intervention Development](#) - Developing our ideas for theoretically impactful interventions

We now know about as much as we could hope to before actually trying stuff out in the field. So it's time to start developing our actual intervention ideas. We'll begin by brainstorming all the theoretically impactful interventions there could be (based on our knowledge so far). This can be broken down into four stages - 1) [defining](#) and understanding the problem and its causes, 2) clarifying which factors are malleable and have the greatest [scope for change](#), 3) Identifying how to bring about change: [the change mechanism](#), and 4) identifying [how to deliver](#) the change mechanism(s). After this, we should have a bunch of (ideally 10+) ideas that could theoretically be impactful. We're going to shortlist them to our top 3-5 ideas and test their viability in the field. Before we do that though, we're going to need to have strong monitoring systems, so we'll need to take a quick detour to look into Impact Monitoring.

6. [Impact Monitoring](#) - How we will monitor and evaluate our intervention(s)

We'll first explore [Data Collection](#), specifically using [the CART Principles](#), ensuring the data we collect is a) **Credible**: Collecting high quality data and analysing the data correctly, b) **Actionable**: Committing to acting on the data we collect, c) **Responsible**: Ensuring the benefits of data collection outweigh the costs, and d) **Transportable**: Collecting data that generates knowledge for other programs. Then we'll build our first iteration of our [Monitoring & Evaluation](#) (M&E) system(s), before developing our initial [Cost-Effectiveness Analysis](#) of our potential intervention(s) - which should help us get to that shortlist!

7. [Viability Testing](#) - Testing the viability of our most promising theoretical interventions

It's now time to develop our [Implementation Plan\(s\)](#), outlining how we intend to do small-scale tests (and refinements) on our shortlist of theoretically impactful interventions, in order to ensure they are actually viable in the context we're working in. Depending on our interventions, we may have a number of [potential deliverables](#) that accompany the plan. When we're ready to go, we'll start doing our small-scale tests on the most promising interventions, by conducting our [Formative Evaluation\(s\)](#). These are different to a Pilot as they are much smaller scale, and are designed to be adjusted as they go (you might call

them a “pre-pilot”). We'll be doing our Pilot next as part of Impact Testing our most promising (and viable!) intervention(s).

8. [Impact Testing](#) - Assessing the potential impact of our most promising viable intervention(s)

We now have an intervention (or maybe 2 or 3) that we think is theoretically impactful, and that works in the context of our target populations. Now it's time to collect sufficient evidence of effectiveness of our intervention(s) to justify rigorous evaluation (such as an [RCT](#)), or to start our implementation. The process for this step is actually somewhat similar to Viability Testing. We're going to start by developing our [Implementation Plan\(s\)](#), though these will likely be more detailed and rigorous than the ones we did for our Formative Evaluation(s), outlining any potential deliverables we want to see as outcomes, before finally moving on to implementing our [Pilot\(s\)](#). You'll likely want to revisit the [Impact Monitoring](#) at this stage, to ensure your M&E systems have been updated to reflect our new learnings.

9. [Growth](#) - Investing in rigorous evaluation or implementation of the intervention

Congratulations! You've completed your “Explore” phase. You now have an impactful intervention to either invest in a rigorous evaluation of your intervention (i.e. an [RCT or Quasi-Experimental Design](#)) or to [grow your organisation](#) around. You'll likely be spending this time building up your organisational capacity and updating your [Theory of Change](#) and [Impact Monitoring](#) systems to continuously improve your internal learning (and external accountability).

10. [Scale](#) - Expanding our proven intervention to scale

Now that you've had the time to rigorously evaluate (or demonstrate the effectiveness of) your intervention, you might reach the point where it's time to scale. In which case you need to detail your [scale strategy](#), then develop your [scale model](#), before finally, [implementing your scaled program](#).

[Appendix](#) - Additional helpful information

Such as a list of the different [Programme Development Methodologies](#) that were used to develop the Roadmap. As well as useful tools for [Brainstorming](#) (which you're asked to do several times throughout the Roadmap!)

Introduction

One of the things that surprised me when trying to develop the intervention for Shrimp Welfare Project (SWP), was trying to understand a clear pathway from A (wanting to help shrimps cost-effectively) to Z (having a demonstrably impactful intervention to scale-up).

It was difficult to get a sense of our Knowns, our Known Unknowns and our Unknown Unknowns. There were existing Programme Development Methodologies that we could draw on, but they were often (on their own) insufficient for our purposes, for example:

- They only covered one aspect of the overall Programme, such as how to create a Theory of Change, how to Design an Intervention, or how to do Impact Monitoring or a Process Evaluation
- They presumed some stages were already in place, such as moving from an Evidence Review to a Process Evaluation - without a step to design and implement an Intervention to evaluate
- Written for a context that didn't quite work for us, such as the Global Health and Development space (and therefore were written with assumptions such as having beneficiaries we could interview)

As a result of this, I spent a lot of time reading up on and learning about Programme Development methodologies - piecing a new one together that I thought would make sense for SWP. I want to share this work in case it's useful to other NGOs, in particular for NGOs in contexts similar to ours - i.e. with a low existing evidence base (such as shrimp welfare), working in new contexts (such as with farmers) and where the beneficiaries are unable to communicate with us (shrimps, in our case). I've attempted to map out the A-Z journey to a scaled intervention - I hope it's useful to you 😊

A quick note - Though this Roadmap is written in such a way as to outline the steps from A to B to C, etc. In reality some of these steps may run in parallel (or at least, overlap in some ways). For example, one member of your team may be conducting a [Literature Review](#), while others are preparing to do some on-the-ground [Context Analysis](#) work. In particular, throughout the Roadmap, you'll likely continue to revisit and reflect on your [Theory of Change](#), as well as updating and refining your [Impact Monitoring](#).

1. Key Decisions (and Tools)

This section is largely inspired by (and in some cases, directly taken from) the Charity Entrepreneurship Handbook - "[How to Launch a High-Impact Nonprofit](#)". In fact, it's probably going to read like this whole first section is essentially a pitch for Charity Entrepreneurship. I hope it doesn't come across like that, but a lot of these concepts (and this structure) came from me going through the CE Incubation Program, as well as these being some of the people who have thought about this whole process the most intensely.

At the outset of your path to impact, you will have a number of key decisions to make. I want to present some decision making tools that should help, as well as present some guidance around how you might end up answering some of these key decisions.

Lots more information about these tools can be found in the Handbook linked above (or indeed, by applying to the [Charity Entrepreneurship Incubation Program](#)!). In particular, details on Decision Making Tools can be found in Part II of the Handbook, and Key Decisions can be found in Part III.

1.1. Tools

To start with, let's present some decision making tools that can be used to help you make your key decisions (and ideally, any other decisions you may need to make on your path to impact!)

A great overview of these decision making tools can also be found at this [EAG talk](#) by Joey Savoie.

- **[Multi-Factor Decision-Making](#)** - Creating a set of criteria, assigning weightings for each, and then going through multiple options to generate a score to help you assess the best options. This enables you to combine a large number of factors into a single score, and be transparent about your reasoning. Often created in a spreadsheet. **This one in particular will be used the most throughout these first few key decisions.**
- **[Rationality](#)** - Using the tools of rationality to make decisions more effective, such as understanding biases, Bayesian reasoning, understanding crucial considerations, steelmanning your views, and having epistemic modesty.
- **[Scientific Method](#)** - Understanding the relative strengths of different types of evidence when reviewing the literature, as well as concepts such as randomisation and falsification if you need to undertake primary scientific research.
- **[Effective Altruism](#)** - Using the insights gleaned from the EA community (themselves often derived from Economics and Philosophy) to improve your decisions. Examples include the [ITN Framework](#), [Moral Circle Expansion](#) and [Expected Value](#).
- **[Independent Experts](#)** - Knowing when to reach out to experts, the different types of experts - i.e. Specialists, Domain Experts or Broad Experts), and how to assess their knowledge. *We'll discuss a practical application in more detail in the [Expert Feedback](#) section of the roadmap.*
- **[Task Planning](#)** - The [ability to prioritise](#) and [focus on high-impact tasks](#), [deep work](#), [maintaining healthy habits](#) etc.
- **[Problem-Solving](#)** - Knowing how to solve problems effectively, which can be done by following a series of steps: 1) Determine how important it is, 2) Define the problem, 3) Generate solutions (divergent thinking), 4) Compare solutions (convergent ideation), 5) Implement your solution(s), and 6) Upstream problem-solving. [Steelman Solitaire](#) is another useful tool that falls under the umbrella of problem-solving.

- **Creativity** - More of an art than a science, creativity as a tool refers to things like: mixing together ideas to create something new, knowing how to generate ideas (and answers), and knowing how to find the “wave” of creativity (i.e. knowing under what conditions you're likely to enter [a flow state](#))
- **Long-term Planning** - Being able to generate an overall goal (such as our upcoming section on the [Theory of Change](#)), while also knowing what level of detail you should plan ahead (i.e. a month-to-month plan should be more detailed than a one-year plan, which in turn should be more detailed than a five-year plan).
- **Cost-Effectiveness Analysis (CEA)** - A CEA is basically a calculation aimed at creating a single number representing a benefit (such as a health metric) for a given cost (normally some currency - e.g. “the result of the CEA showed a life could be saved for “\$3,000”. *We discuss CEAs in much more detail in the [Impact Monitoring](#) section of the roadmap.*
- **Monitoring and Evaluation (M&E)** - Monitoring the right metrics and thoughtfully evaluating the insights that can be gleaned. The concepts of M&E will come up a few times throughout this roadmap, so I won't go into too much detail here. In particular Monitoring is discussed in the [Right-Fit Evidence](#) section and Evaluation is discussed in the [Process Evaluation](#) section. *We discuss M&E in much more detail in the [Impact Monitoring](#) section of the roadmap.*
- **Meta tools** - This refers to concepts such as:
 - [The Explore/Exploit tradeoff](#) - Knowing (or committing to) how long you should spend “exploring” options, before committing to “exploit” the knowledge you have gained. In [Algorithms to Live By](#), this tradeoff is put at 37% spent exploring, the rest exploiting.
 - [Time-Capping](#) - Fixing the number of hours for a certain task, research project or decision and keeping research within those bounds.
 - [Iterative Depth](#) - A process of narrowing down many ideas to the most promising ones, without spending too much time on the least promising ones. This can be achieved by “[time-capping](#)” each stage of the process, and iteratively researching ideas that make it to subsequent stages in increasing depth.
 - [Narrow Focus](#) - The ability to maintain focus on an idea without spreading yourself too thin. Ten half-finished projects are likely not as valuable as a single finished project.
 - [Feedback](#) - Giving and receiving feedback, as well as knowing which pieces of feedback to implement and which to discard.
 - [Reevaluation Points](#) - Building in regular times that you'll reevaluate your project with fresh eyes and perspective, allowing you to maintain [narrow focus](#) at other times.
 - **Components of a Good Tool** - Understanding where a tool you might use falls across the spectrum of: Speed, Cross-Applicability and Accuracy.

- **Tool Trade-offs** - Choosing a tool based on the decision, such as between a) a fast, but semi-accurate tool, or b) a slow but accurate one. Often depends on the importance of the answer, how much time you have, and number of options being compared.
- **Using Multiple Tools** - Often multiple tools can be used to get closer to the truth, and if multiple tools are converging on an answer - you're likely heading in the right direction!

1.2. Values

A good resource on this can be found within the [Know Your Values](#) article on the [Good Enough](#) website (again, I'm linking to content created by Joey Savoie).

You may already have a good sense of this, but I found the process of working through this and trying to concretely articulate my values to be a worthwhile process. Being able to articulate your values will help with the rest of the decisions in this section.

A useful tool for this is [the Moral Parliament](#), which is a practical way of dealing with your own Moral Uncertainty. Rather than going all-in on a specific moral theory, pretend you have an imaginary parliament of 100 seats to fill. How many of each of these seats would you assign to different moral views? You can keep this fairly high-level (like 80% utilitarianism, 15% deontologism and 5% virtue ethics), or ideally, pretty specific (i.e. 15% Negative Utilitarian, 15% Total Utilitarian, 10% Average Utilitarian, 10% Hedonistic Utilitarian, 10% Preference Utilitarian, 10% Rule Utilitarian, 10% Motive Utilitarian, 15% Deontologist, 5% Virtue Ethicist, etc.). These are made-up numbers but I hope you get the idea.

You can then work through different questions within your parliament, figuring out how they would vote on different moral questions (potentially tweaking your percentage of seats as you go, if you don't believe your current parliament accurately affects your views).

1.3. Career Area

Your best bet here is probably going to be [80,000 Hours](#) Career Planning tools. I really don't want to reinvent the wheel here, so although this is likely the most important decision in this whole process, it's also one of the shortest sections in the Roadmap.

The first of your big key decisions now is going to be your career path. Is founding a charity the best way to go? Or does your personal fit and values suggest a different career path is optimal?

Even if you're pretty set on the idea of creating an idea and seeing it through to fruition (as is the intention of this Roadmap), Non-profit Entrepreneurship might not be the best fit for you

(or your idea), and instead For-Profit Entrepreneurship, or Social Entrepreneurship (a double-bottom line) might be a better choice. Additionally there are alternatives, such as working at an existing [Effective Altruism organisation](#), or in [Research](#), or [Communications](#), or [Policy](#), or [Earning-to-Give](#), or developing [Career Capital](#).

Another good video on this (again, by Joey Savoie) is this talk and Q&A by Charity Entrepreneurship - "[Impactful opportunities around and adjacent to charity entrepreneurship](#)"

For me, this looked like using the [80,000 Hours framework](#), creating a Weighted Factor Model (I used the 80,000 Hours "[Dream Job](#)" criteria), and assigning each criteria a "weight" according to how important each factor was (with Personal Fit being a multiplier to the overall score), and then assessing a bunch of different Career Paths against this, for example:

- **Engaging Work (20%)** - Is the work important and engaging? (1 = work that is not important or interesting; 5 = work that is highly engaging and potentially impactful)
- **Work that helps others (35%)** - How much good will I be doing? (1 = Harmful and/or waste of money; 5 = small improvement of top charity or comparable)"
- **Supportive Colleagues (15%)** - Can I get help/social support from colleagues when I run into problems? (1 = unsupportive colleagues, hostile work environment; 5 = very supportive colleagues, learning atmosphere, "disagreeable givers")
- **Basic Needs (10%)** - How much will this benefit me? (1 = Long commute, long hours, unfair pay, job insecurity; 5 = Short/no commute, work-life balance, adequately paid, job security)
- **Career Capital (20%)** - How much will this benefit my career? (1 = Wouldn't put it on my CV; 5 = Significantly enhances CV and/or good networking opportunities)
- **Personal Fit (force multiplier)** - How much better suited am I to this role than the next best candidate? (0.5 = Most people can do this as well or better than me; 2 = I can do this job better than most people I can think of)

A final score was then calculated for each career option, by adding up the total score I'd given for each criteria, multiplied by the respective weighting of that criteria (with Personal Fit being the exception, as I multiplied the total score of all the previous criteria by personal fit, to give my final score).

1.4. Charity Idea

I think this next step would also make sense if your Career Area wasn't to found a charity or a new intervention, such as working for an established organisation, earning-to-give, or other forms of entrepreneurship. But I'm not certain of this, and the rest of this roadmap is written with the assumption that you would be starting a charity, so I'm going to keep the language as such.

Next up, you'll need to decide on a (charity) idea. You should be able to brainstorm ideas within a Cause Area based on your defined [Values](#) above, as well as brainstorming ideas that fit within your chosen [Career Area](#).

If you don't have this already, you'll likely need to go through a process to determine this (or use a pre-vetted cause area, such as those identified by [Charity Entrepreneurship](#) or [CEARCH](#)). If you want to go through a similar process yourself, Peter Wildeford breaks down how this decision can be made for the question of "[What Charity to Start?](#)" using the Multi-Factor Decision-Making tool:

1. **Come up with a well-defined goal**
2. **[Brainstorm](#) many plausible solutions to achieve that goal**
3. **Create criteria through which you will evaluate those solutions**
4. **Create custom weights for the criteria**
5. **Quickly use intuition to prioritise the solutions on the criteria so far**
6. **Come up with research questions that would help you determine how well each solution fits the criteria**
7. **Use the research questions to do shallow research into the top ideas**
8. **Use research to re rate and rank the solutions**
9. **Pick the top ideas worth testing and do deeper research or MVP testing, as is applicable**
10. **Repeat steps 8 and 9 until sufficiently confident in a decision**

Another good resource I'd recommend that discusses this concept is "[How to do Research that Matters](#)" by Karolina Sarek.

1.5. Co-Founder

A good overview of this process can be found in the CE article "[How to Successfully Pick a Co-Founder](#)"

Two co-founders is often the sweet spot. You likely want to find someone who has complementary skills and psychology to you, as well as shared [Values](#) (hopefully your articulated Values above can help to ensure you're on the same page with your potential co-founder). Additionally you'll want to find someone who you personally have a strong personal fit with, to the end that collaboration with this person produces better work than either of you could have produced independently.

The best way to know if you work well with someone is to work with them. So ideally you'll want to work on some discrete projects with different people, to test your fit with them. Finding a co-founder might be tough, but the [Charity Entrepreneurship Incubation Program](#) is designed to help you test your fit with a number of different co-founders, but the basic

concepts should apply outside of the context of an Incubation Program like this (I also promise I'll stop talking about CE and the Incubation Programme after this first section, but they really are the experts at this first stage of the Roadmap).

After working with a bunch of different talented people, and getting a sense of who I thought was a good fit, the deciding factor for me was filling in a "Founders' Agreement" together (*this idea is outlined in more detail in the CE article ["How to Strengthen Your Co-Founder Relationship"](#)*). This really helped solidify my understanding of what the actual relationship would be like, before committing to starting the project together. Effectively this looks like working through a document together that looks at the following:

- **Roles and Responsibilities** - Who has overall ownership of different aspects of the organisation. For example, which of you has ownership of the following aspects of the charity (or is this shared):
 - Strategy (Planning / Fundraising)
 - Outreach (Advisors / Stakeholders)
 - Operations (Technical / Legal / Financial)
 - Human Resources (Management / Recruitment / Internal Culture)
 - Communications (Written / Visual / Verbal)
 - Research (Monitoring & Evaluation / Cost-Effectiveness Analyses)
 - Etc. etc.
- **Decision-Making** - How are decisions made at the organisation? Does this change if it's an "important" decision? How are we defining "important"? Do we sense-check decisions with external mentors? How often?
- **Field-Time** - What's the expectations of co-founders spending time in-country, learning from your beneficiaries. From an organisational perspective, leadership time in-country is incredibly important, but from a personal perspective, there might be limits. It's good to consider this and set expectations up front.
- **Work/Life Balance** - What are the expectations for work? Do you prefer working 10 hours a day, but 4 days a week? Are there other considerations? How will we prevent burnout? Do we have any policies like not working weekends? Or committing to a certain amount of vacation?
- **Communications** - What communication channels do we use, and what are the expectations for each (i.e. non-urgent topics are on Slack, urgent are on WhatsApp. Or perhaps WhatsApp is off limits for work, but a phone call is instead preferred for urgent matters, but only within a certain timeframe, etc).
- **Personal Happiness and Growth** - Longer term personal development (or even burnout) should not fall victim to daily operational tasks. Outline the areas where the co-founders commit to taking care of themselves (i.e. through exercise or meditation etc) and where each co-founder plans to develop (these could build on strengths, or tackle weaknesses - depending on organisational priorities).

For me, a lightbulb moment was going through the roles and responsibilities with different potential co-founders, and realising that with Andrés (now my co-founder of SWP), the stars really aligned - the responsibilities I wanted to take on, were ones he didn't want to take on, and vice versa. It really solidified to me that our skill sets were complementary and we'd make a good team.

1.6. Country Selection

We're now at the final step of our Key Decisions - it's time to select (or narrow down) your implementation country. This is not set-in-stone, though bear in mind that once a location is selected, it can be difficult to change, and location-specific knowledge and skills that you invest in building may not be directly transferable to other contexts.

The process of determining a country is both practically decision-relevant (it gives you a place to start), as well as helping you think through the factors to consider for selecting an implementation country (such as scale, neglectedness and tractability for the problem in that country - what are the proxies you'll be using to assess that, and how will you answer these questions).

Again, we're going to fall to our trusty Weighted Factor Model to help us make this decision, but in this case many of the answers to your criteria won't be intuitions, but concrete data (for example, for us, "Shrimp Production in Thousand Metric Tonnes" was a datapoint we could reliably find across different countries, and is a useful proxy for Scale).

You'll want your Country Selection model to factor considerations such as the following:

- **Addressable Need** - This could vary significantly depending on your [Charity Idea](#)
 - **Scale** - What is the scale of the problem in country (we could use proxies like "shrimp production tonnage", or "% of shrimps exported to the EU")
 - **Neglectedness** - Is anyone else working on this problem? And/Or - "How bad is the problem" (for example, we could use proxies like "Average Water Quality" based on market research)
- **Tractability / Feasibility** - To what extent are problems theoretically solvable?
 - **General Infrastructure** - You might want to also consider things like access to electricity or phones, depending on your Charity Idea
 - **Proximity to talent** - Are you able to hire talented, dedicated people in this country (or willing to move to this country)
 - **Traction** - Is it possible to get traction on this issue in this country? (If it's particularly difficult to work on a certain cause area in a certain country, then even if the problem is large and neglected, it may not be worth pursuing)

- **Corruption / Bureaucracy** - Is corruption, or excessive bureaucracy a limiting factor in this context? (There is data such as the [Corruption Perceptions Index](#), which can help here)
- **COVID-19 (or similar)** - This was a big factor for charities founding around this time, you may also end up having to factor in considerations like this (i.e. travel restrictions, or vaccination rates)
- **Personal Fit** - This will likely be [Co-Founder](#) specific
 - **Global Peace Index / Democracy Index** - An index such as these (or both) may help you determine personal fit
 - **Affordability** - How affordable is it to spend significant time in the country? (You can also use something like the Cost of Living Index to help with this)
 - **Willingness to spend significant time** - How much is each co-founder willing to commit to being in this country?

2. Theory of Change

A Theory of Change (ToC) outlines the logic behind a program and explains how it is expected to bring about impact. It demonstrates how the program's activities and use of resources will lead to changes in behaviour and improvements in the lives of people (or, in our case, shrimp). This theory is an important foundation for any development program.

It's important to note here (and this is probably the single biggest difference between the discussion of a Theory of Change in this Roadmap, than with other [Programme Development Methodologies](#)), that at this stage the Impact Roadmap assumes that you don't have a detailed plan ready-to-go, or an in-depth understanding of the problem you're trying to solve. Our first iteration of our Theory of Change is effectively going to be a sketch, where we work backwards from "Impact" and try to detail the "Outcomes" that need to happen in order for the Impact to happen. And then work back from there to get an idea of the Activities that will need to take place to get us there (we'll define these terms in more detail soon). The larger important point here is that our Theory of Change will be a "first attempt" at mapping change, we're going to be using the rest of the Roadmap to develop our understanding of our path to impact, so we can reflect that in subsequent iterations of our Theory of Change.

To reiterate then - **The Theory of Change should be referred to (and if needed, updated) often.** It should function as a living document. Ideally, **it should be returned to after every subsequent (major) stage in this roadmap** (i.e. [Evidence Review](#), [Needs Assessment](#), [Intervention Development](#) etc.) as at every stage you'll be learning more and more, and those insights can and should be reflected in your [Theory of Change](#).

2.1. Project Objectives

As outlined by the [Social Impact Navigator](#), it helps to have clear Project Objectives, to orient your projects goals, in order to help you develop your Theory of Change diagram.

Project Objectives detail a project's intended results, which contribute to improving A, B, C, for X, Y, Z. Developing them allows you to orient yourself (and the team) to ensure you're on track, motivate your team, demonstrate your work (to external stakeholders, and funders), and aids with your [Impact Monitoring](#) down the line. They become the foundation for your project, and ultimately will likely form the "Impact" section of your Theory of Change diagram, which you can then work back from, to get a more grounded understanding of your project's Activities. As with the Theory of Change diagram, the Project Objectives will likely not remain stable over time, and should be regularly revisited, and updated if required.

There are two levels of Objectives and it might make sense to define project objectives for only one of them:

- The **Target-Group Level** (if the problem were to be **solved**, what would the situation look like?)
- And the **Societal Level** (if the problem was **eliminated**, what would the situation look like?)

Additionally, SMART criteria are helpful in formulating good project objectives:

- **Specific:** Try to be as precise and transparent as possible when defining objectives, thereby enabling third parties to understand them, too.
- **Measurable:** Objectives must be subject to quantifiable description.
- **Accepted:** Project objectives need to be accepted by stakeholders. It is important to establish a common understanding of objectives that is accepted by all parties involved.
- **Realistic:** It must be possible to attain the formulated project objectives. You do not have to be 100% sure that objectives will be met, but they should be doable.
- **Time-bound:** You should be able to provide an estimated schedule for achieving project objectives. Objectives can be reached during the project's operation or later, the main thing is to have a schedule to keep the project on track.

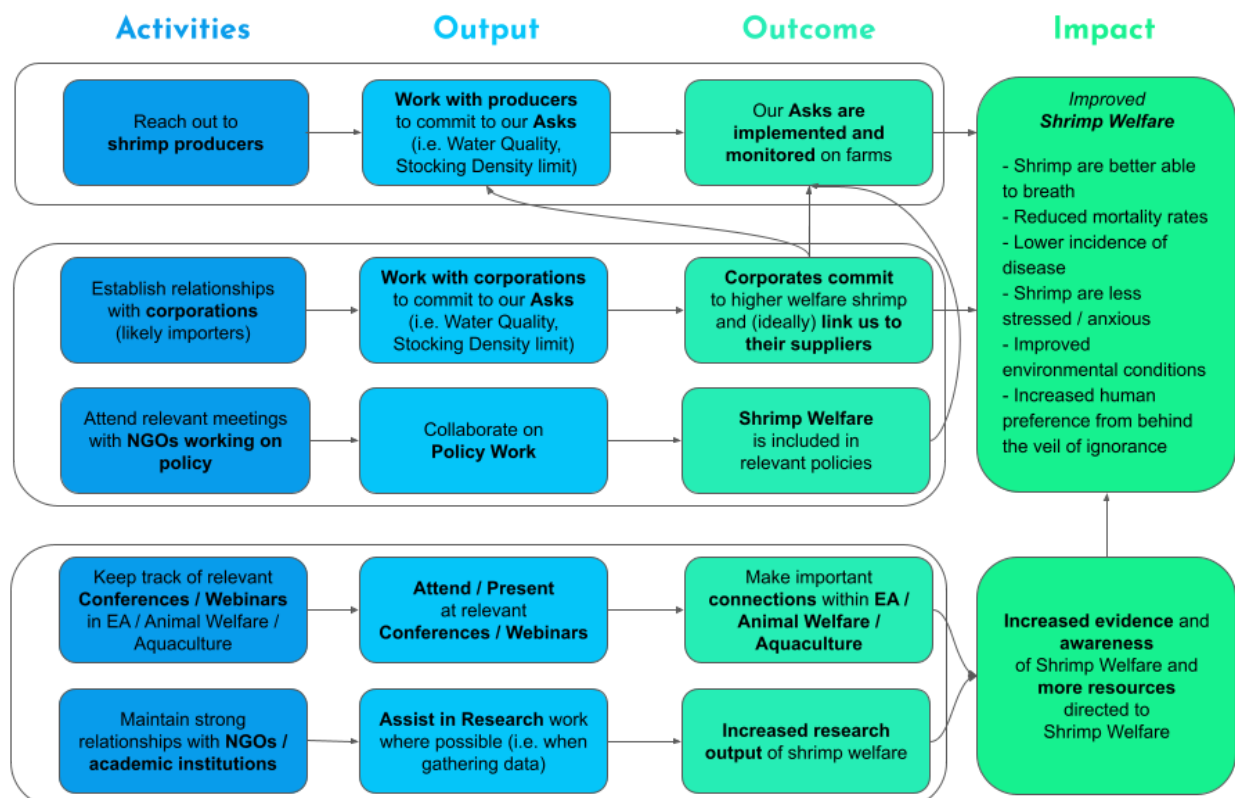
One final important point to emphasise here - **Programme Activities are not Project Objectives**. What this means is that the Activities your project carries out, is not the endline goal of your work (the Objective). Your Activities start a chain of events that lead to the projects Outputs, which lead to Outcomes, which leads to your Impact (and in so doing, achieves your Objectives). **"Distributing bednets" is an Activity. "Reducing the prevalence of malaria" is the Objective.**

2.2. High-level Theory of Change

A diagram outlining the process of change

The Theory of Change diagram links (visually) how we expect to achieve Impact through our Activities (based on theory and evidence, such as: academic evidence, lived or professional experience, and sound reasoning)

- **Activities** - The concrete set of inputs and activities we carry out (i.e. reach out to shrimp farmers and convince them to commit to our Standards)
- **Outputs** - The expected result of those inputs (i.e. shrimp farmers implement our standards on their shrimp farms)
- **Outcomes** - The changes we should see as a result of the outputs (i.e. water quality etc. improves on shrimp farms we work with)
- **Impact** - The impact that will lead from the outcomes (shrimp welfare is improved on the farms we work with)



There are a number of different ways to do Theories of Change, and they can be as complex or as simple as you need them to be. Often it makes sense to produce a simple, clear Theory of Change, accompanied by a Narrative, to share externally (i.e. with funders), in

addition to a more complex, less-polished but detailed Theory of Change, in order to tease out assumptions and get internal synthesis among the team.

Although Theories of Change are often read left to right, it often makes sense to create them right to left. As in "for Y to change, we need X as a precondition", which is why it often makes sense to start with defining your [Project Objectives](#).

2.2.1. Identifying Assumptions

You're not finished when you've drawn out your diagram, there's still one last step, which is to identify the assumptions that your Theory of Change rests on, such as:

- **Causal** - Assumptions about the causal mechanisms for change
- **Contextual / Environmental** - Assumptions about your operating environment
- **Normative** - Assumptions about what 'ought' to be

It also makes sense to try and track the links between every box and node in the process, as well as to really articulate them in detail. For example, you might label every step in your Theory with a number, which can then be referenced in a table/spreadsheet which can allow you to write more detail and clarify points about each step, without cluttering up the diagram. For example, we created a table with the following headings:

- **#** - Which step(s) in the Programme Theory does this relate to?
- **Note / Risk / Assumption / Externality** - Is this a general **Note** about the step? Or a **Risk** (something that could go wrong, likely pulling from our Crucial Consideration)? Or an **Assumption** (an underlying assumption that we want to surface)? Or an **Externality** (or a potential positive or negative result that may occur)
- **Description** - What is the Risk / Assumption / Externality we have identified
- **Likelihood** - What is our expected likelihood of this happening (we used a %)
- **Impact** - What is the potential impact if this happens (we used Small / Medium / Large)
- **Mitigating Actions** - What mitigating actions can we take to prevent this from happening (or what will we do if this happens)

Break down every assumption to be as concise as it can be, such as breaking down '**Users will pay for this model**', into '**Users are interested in the model**' and '**Users will have the means to pay for it**'.

3. Evidence Review

An evidence review is a comprehensive examination of existing research on a specific program. It helps to understand what is known about the issues the program aims to

address and its design. An evidence review can help answer questions such as: **What interventions have been tested or studied in relation to this program? Under what conditions were they implemented? Were they effective? Why or why not?** Conducting an evidence review can provide valuable insights into the program and inform its development and implementation.

3.1. Literature Review(s)

This step of the process largely follows the [Evidence Review section of the ID Insight Guide](#).

A literature review is a systematic review of the relevant sources to look for evidence related to your program. The primary benefit of this is to have decision-relevant data for developing your intervention, but in many cases it makes sense to also write up and publish the literature review so that it can be shared with interested parties.

It's worth noting that in some cases, you may be able to find a high-quality literature review compiled by someone else. If you're confident in someone else's work, this step can be skipped.

1. **Annotated Bibliography** - The first step is to scour the literature for evidence that might be decision relevant (these shouldn't just be academic, but should also include things like public datasets, or websites of organisations working in this space), and compiling these references into a bibliography in a single place (like Google Sheets, or Notion, or Mendeley). During this process, it helps to add key takeaways (or quick write-ups) for each source, creating an *annotated bibliography* - in particular this helps if you have multiple people working on this, as it can help prevent duplication of effort.
2. **Weighting Evidence** - Then you need to weight the quality, generalisability and usefulness of the evidence you've gathered, likely using a [Weighted Factor Model](#):
 - a. **Quality** - Assessing the quality of this evidence, using criteria such as:
 - i. Risk of bias
 - ii. Consistency of effect
 - iii. Imprecision
 - iv. Indirectness
 - v. Publication bias
 - b. **Generalisability** - Is it possible to use evidence from another context to this program, J-PAL recommends a four-step framework to understand this:
 - i. What is the disaggregated theory behind the program?
 - ii. Do the local conditions hold for that theory to apply?
 - iii. How strong is the evidence for the required general behavioural change?
 - iv. What is the evidence that the implementation process can be carried out well?

- c. **Usefulness** - Finally, try to assess whether the evidence enables you to understand:
 - i. The scope of the problem
 - ii. The major constraints to solving the problem
 - iii. The success of previous attempts to solve the problem
 - iv. The links and assumptions in your [Theory of Change](#)
3. **Write-up** - Finally, you can write-up your findings from previous studies (highlighting the relevance of evidence to your program, as well as their rigour) as well as any remaining evidence gaps (particularly links or [assumptions](#) in your Theory of Change).

3.2. Expert Feedback

The Charity Entrepreneurship Handbook - "[How to Launch a High-Impact Nonprofit](#)" has a whole chapter on Independent Experts, which I'd recommend.

Depending on your intervention, this might not be a necessary (or viable) step of your Evidence Review, but it's likely that it will. You probably have some preliminary conclusions around what your intervention might look like at this point, and you're about to move on to conducting a Needs Assessment. Before you do, it often makes sense to gather some Expert Feedback, likely based on the results of your Evidence Review so far (as a sense-check - Have you missed any key pieces of evidence? Have you interpreted your conclusions correctly? What information would they choose to gather as part of a Needs Assessment?).

Experts can provide synthesised analysis and a broader scope, often able to directly compare strategies in a broader context, and give you a sense of the consensus in your chosen area. They're often slow to update their views, which can provide something of a safety net. They can be a great source of missed information, can find weaknesses and apply common sense filters.

Additionally, it probably makes sense to reach out and develop relationships with experts at this time. You'll be fairly well-versed in the topic as you'll have just completed a literature review, so you'll be using their time respectfully, and your questions will likely be thoughtful, enabling their answers to be insightful. But you've still got a long way to go before you have a working intervention, and having a strong network of experts you can consult with from here on is going to be a real help.

As we alluded to in our quick discussion on experts during the [Tools](#) section of this Roadmap, there are three types of experts:

1. **Specialists** - Narrow knowledge about something very specific (e.g. shrimp disease)

2. **Domain Expert** - broad knowledge about a narrow domain (e.g. shrimp farming / aquaculture)
3. **Broad Expert** - Broad, comparative knowledge (e.g. evaluators of animal welfare interventions)

It's worth noting however that you should exercise some caution around expert feedback. Experts are people too, they're fallible and prone to bias, they may have reasons to not be fully transparent with you, and it can be too easy to generalise and cross-apply their advice. Additionally, they can be limited in how specific/decisive they're willing to be, often sceptical of providing numerical estimates.

So how do we go about getting Expert Feedback? This is often more tractable (and ultimately less intimidating) than it at first seems. Often experts are excited in talking to people interested in their field of expertise (particularly for example, if it's an academic in a niche field). You should make sure to prepare before contacting an expert, demonstrating that you're aware of the value of their time, and prepare specific questions in advance. Make it clear that you respect their time, what you're intending to use the information for (i.e. your transparency policy, is this for internal decision-making only? Do you intend to write a report quoting them? etc), that you'll work around their schedule, and you'll be brief and specific. Often you might find that meetings overrun and you get more time with the expert than you bargained for, which is probably a good thing! Just don't jump out the gate asking for hour long meetings, if a 15-30 minute meeting would likely suffice.

This process can be somewhat systematised. You can set a goal to contact X number of experts, writing a short, concise email explaining your goals. If possible, have a set of systemised questions you can ask a number of experts and then quantify the answers (you may also have specific questions that only 1 or 2 experts can answer). From here, you can synthesise answers (likely in a spreadsheet).

4. Needs Assessment

To be effective, a program must address the needs of the community it serves. Gathering input and perspectives from the community where the program will be implemented allows for a thorough understanding of the problem the program aims to solve. A needs assessment is a tool that helps identify the challenges facing the target population and the broader economic, political, and social context in which the program will operate. Incorporating a needs assessment into program planning can help ensure that the program is responsive to the needs of the community.

I want to make a quick note here about the placement of the Needs Assessment within this Roadmap, as it's another point where the Roadmap differs from other [Programme](#)

[Development Methodologies](#) (for example, PHINEO's [Social Impact Navigator](#) places this as step 1, before defining your Theory of Change, and IDInsight's [Impact Measurement Guide](#) has this step take place after the ToC but before the Evidence Review). However, we found that it made the most sense to first develop an idea for how we want to make an impact (the [Theory of Change](#)) first see what the existing literature has to say (the [Evidence Review](#)) before trying to do some in-person work (the [Needs Assessment](#)). Another reason to do the Needs Assessment after the first two stages is to prevent potentially replicating work that has already been done (which would be uncovered during an Evidence Review), or potentially jeopardising beneficiary/stakeholder relationships by coming off as unknowledgeable. As noted elsewhere in the Roadmap however, there may be some overlap between sections within the Roadmap, so doing some Needs Assessment work while the Evidence Review is ongoing, for example, might be the most effective use of time and resources.

4.1. Context Analysis

This step of the process largely follows the [Needs Assessment section of the ID Insight Guide](#).

Much like the primary output of the Evidence Review stage is the production of a Literature Review. The primary output of the Needs Assessment stage is likely going to be Context Analysis (likely written up into a Report). The goal is to get a much stronger understanding of the context in which you plan to work

Effectively you're trying to get answers to some key questions in the context you're trying to work in:

1. **What do we want to learn?** - You will undoubtedly have unanswered questions now, some will have arisen during your Theory of Change development, others during your Evidence Review. Either way, it's time to go into the field. What questions do you want to have answered? (At this point it's worth just thinking about the questions you actually want to answer, you might end up asking the actual questions differently in the field, but the first step is to understand what you actually want to learn).
2. **What assumptions are in our Theory of Change?** - Next up, let's really zoom in on our Theory of Change (to the extent that you're able to at this point). What are the assumptions that are buried there? Really tease them out, and make them as distinct and clear as possible (the heavy lifting here should already have been done during the development of the Theory of Change within your [Identifying Assumptions](#), but it's worth revisiting and potentially updating here)
3. **How well do our assumptions need to hold?** - Which of these assumptions are the most important? Which ones are unimportant? It might make sense to prioritise with a matrix grid at this point, assessing your assumptions against your certainty and risk for each (prioritising data collection that sheds light on assumptions that seem likely

and would have the most impact in your Theory of Change - as a reminder, this data is also likely in your [Identifying Assumptions](#) work, it's all coming together nicely 😊).

4. **What data do we need?** - We're going to go into this in a lot more detail in our [Impact Monitoring](#) section, depending on how robust your data needs to be at this stage, it may make sense to skip ahead to that, to get some tips on collecting high-quality data.

From there, you can conduct the Needs Assessment - we'll go into more detail on executing an on-the-ground project like this as part of the [Viability Testing](#) stage of the Roadmap, but I don't want to leave you high and dry at this stage... So if you need some help in the planning and executing of a data collection process like this (and don't mind some light spoilers!) Please feel free to read the [Implementation Plan](#) section. The Needs Assessment process will then follow these four broad steps:

1. **Collect data** - Again, in our Impact Monitoring section there is advice on [data collection](#), so I'll link to that here in case it's useful.
2. **Analyse the data** - And final time 😊, in our Impact Monitoring section there is advice on [data analysis](#), so I'll link to that here in case it's useful.
3. **Write-up** - In a Scoping Report (largely for the benefit of researchers and funders) and
4. **Update** - Your Theory of Change, or add any new areas of research for your Evidence Review

4.2. Primary Research

Finally, based on the outputs of both your [literature review](#) and your [context analysis](#), you likely still have some unanswered questions. In this case, it may make sense to pursue some primary research before (or while) you develop your intervention. This may involve you taking on the research within the organisation, or trying to realise the research some other way (i.e. by outsourcing it to a research organisation or a university). The way you'll need to approach this will largely depend on the type of research that needs to be done. Therefore this section will have the least actionable guidance, but I'll give some examples of what this looked like for Shrimp Welfare Project.

Primarily for us, this resulted in a project we called "The Welfare Gap" where we used a [Weighted Factor model](#) (hello again, friend) to rank the priority of work we wanted to see, then set about trying to close the gap. We'd originally envisioned this as a number of research proposals, that we would then network with academics and funders. But it ended up being somewhat more diverse than that, for example:

- There were some projects, like a Shrimp Welfare Index, that were already being developed, so we lent a hand in shaping the direction of that project and testing it in a real world context
- Some work was primary research we were able to attempt on a small scale on our own - we wanted to get a sense of consumer awareness of shrimp welfare issues, and whether awareness of these issues would increase their Willingness to Pay (WTP). We were able to use Prolific to test this on a small scale and get some promising results.
- And finally there were some projects that were more traditional academic pieces of work, but in some cases had already been attempted and dropped (such as research on shrimp stunning before slaughter). So in those cases, we worked to understand the reasons and then unblock the work (by facilitating connections between academia, funders and industry).

5. Intervention Development

This stage follows the first 4 steps outlined in the [6SQuID methodology](#) (steps 5 and 6 of the 6SQuID methodology are the [Viability Testing](#) and [Impact Testing](#) sections of the Impact Roadmap, respectively).

As outlined by 6SQuID - once a problem has been identified as needing intervention, the process of designing an intervention can be broken down into six crucial steps: (1) defining and understanding the problem and its causes; (2) identifying which causal or contextual factors are modifiable: which have the greatest scope for change and who would benefit most; (3) deciding on the mechanisms of change; (4) clarifying how these will be delivered; (5) testing and adapting the intervention; and (6) collecting sufficient evidence of effectiveness to proceed to a rigorous evaluation.

I found the 6SQuID methodology tremendously useful in developing our intervention - however it took a little while for me to fully comprehend how much effort needed to go into each step - and what each step might look like. So for each step within 6SQuID, I've tried to break them down further (along with templates) to give a sense of what working through these steps might look like in practice. As a result, this ends up being one of the longest sections in the Roadmap, and writing it out surprised me with how many sub-sections there were to each section.

5.1. Problem Definition

6SQuID step 1: Define and understand the problem and its causes.

Based on the previous stages of the roadmap, we're confident that we've identified there is a problem that requires an intervention. The next step is to *clarify* the problem with stakeholders and existing research evidence. This part seemed simple enough to me on face value, but actually working through it effectively results in **5 key outputs(!)**

In all cases here, you need to create some sort of shareable grid of information - we used Notion Databases, but you could likely also use spreadsheets or documents (or other tools) just as well.

5.1.1. Definitions

Clear and detailed terms to avoid ambiguity and confusion.

Now that you've gone through the [Evidence Review](#) and [Needs Assessment](#) stages, you've presumably learned a whole lot about a very niche topic. It's time to catalogue your definitions somewhere to add clarity to the remaining steps in the intervention development (this is also a useful process for "bus-proofing" the organisation, as well as onboarding new staff).

We used the following Headings (with an example given for each):

- **Concept** - Eyestalk Ablation
- **Area** - Welfare Concerns
 - You may not need this heading, but we found it useful to group together similar concepts, in areas like 'Pond Types', or 'The Shrimp Supply Chain', or 'Sentience'
- **Definition** - Eyestalk ablation is the practice of slicing, squeezing or cutting off an eyestalk to make them reproduce faster. *Shrimps have a gland behind their eyes which tell their ovaries when to mature (the Gonad Inhibiting Hormone - GIH). Without this inhibiting hormone, a shrimp's reproduction system is sent into overdrive.*
 - The definition should be short and clear - however in some cases like this, you may need to provide some more context along with the definition itself.
- **Position** - This practice leads to suffering shrimp and proliferating disease. The offspring of non-ablated shrimps have demonstrated increased resistance to disease.
 - We found it useful to include an official SWP position on concepts where applicable - this meant that it was easy for anyone in the organisation to understand where we stand on potentially controversial topics.
- **Assumption(s)** - Faster reproduction shouldn't be the goal of hatcheries for their broodstock - more resilient offspring results in more sustainable shrimp production.
 - In some cases, we also included any assumptions baked into our definition or position. This was another attempt (like the Theory of Change) to surface underlying assumptions that, if not checked, could mislead us.

5.1.3. Causal Pathways

What shapes and perpetuates the problem

Having defined the problem, we need to understand the immediate (proximal) and underlying (distal) influences that give rise to it. It is only by understanding what shapes and perpetuates the problem (the **causal pathways**) that we can identify ways to intervene.

[Need to give an example](#)

5.1.5. Current Standards

What interventions/policies exist and why they are inadequate

This may look very different depending on your cause area. For us, this meant compiling all the Standards we could find such as standards of Certification schemes (such as ASC or BAP), any standards in place by governments (i.e. policies), standards set by specific corporations (often on the demand side), or standards in place at the producer level (i.e. a farmers Standard Operating Procedures)

This part should absolutely be tailored to your needs in collecting this data (and you may have a lot of it from your Evidence Review), but I'm hopeful that sharing our template is useful for you. We used the following Headings (with an example for each):

- **Standard** - Provisional maximum for stocking density shall be set 15 post larvae/m² (Semi-Intensive)
 - The actual standard provided by the source
- **Welfare Concern** - Stocking Density levels
 - The area of welfare the Standard relates to
- **Domain** - Space Requirements
 - You may not need this, but we wanted a way to group similar welfare concerns - for example we had domains like:
 - Water Quality
 - Space Requirements
 - Stunning and Slaughter
 - Health and Disease Management
 - Enriched Environment
- **Source** - [Naturland Standards - Organic Aquaculture](#)
 - (the title and link to the source of the standard)
- **Notes / Comments** - The lowest stocking density standard we found for a certifier. Though as it's an organic certification it likely isn't as scalable as others.
 - Any notes or comments we have about the standard

5.2. Scope for Change

6SQuID step 2: Clarify which factors are malleable and have the greatest scope for change

Which of the factors that shape the [Problem](#) have the greatest scope to be changed - this could be at any point along the [Causal Pathway](#).

Most interventions take place within systems and exert their influence by changing relationships, displacing existing activities and redistributing and transforming resources within. Some stakeholders may only change their behaviour if other stakeholders do so too (i.e. farmers may only improve welfare standards if they are required to by a certifier, and they need to be certified in order to sell to a certain market). Interventions that address complex problems through multilevel actions are more likely to maximise synergy and long-term success.

5.2.1. Causal Factors

Identify the factors malleable to change, and assess their importance and tractability

At this point, we're essentially just trying to understand the malleability of each link in our [Causal Pathway](#). We'll move on to figure out the [Change Mechanisms](#) and [Intervention Design](#) in the next two steps. It probably makes sense to work through all of the Causal Factors identified in the Causal Pathway, just to be thorough - though there might be some where it is obvious no scope for change, in which case you may choose to skip those.

We used the following Headings:

- **Causal Factor** - Which causal factor are we analysing?
- **Primary / Secondary** - Does this factor impact secondary prevention (i.e. improving the welfare of shrimps on farms) or primary prevention (i.e. preventing shrimps from being brought into existence in the first place), or both?
- **Structural / Proximal** - Is this factor an "upstream" structural factor (which typically take longer and are more challenging to modify, but if achieved have the greatest [population](#) impact). Or is it a "downstream" proximal factor (which is often quicker and easier, but affects a smaller population)?
- **Evidence willing to change** - What evidence do we have that this factor is malleable to change (likely drawing on the [Evidence Review](#) and [Needs Assessment](#) work)?
- **Systems impacted?** - Which systems are likely to interact with the intervention (e.g. if we decide to focus on just producers, which other steps in the supply chain are impacted?)

- **Other systems** - Which other systems need to be/can be modified as well (e.g. do we - or other NGOs - need to convince corporations to require higher welfare standards first in order to require producers to implement our standards? Are they willing to do that?)

5.3. Change Mechanisms

6SQuID step 3: Identify how to bring about change: the change mechanism

Once the most promising (and modifiable) [Causal Factors](#) have been identified, how do we achieve this change? What are the Change Mechanisms we can use (i.e. the process that triggers change for individuals, groups or communities)?

5.3.1. Populations

Who is currently most/least likely to benefit from an intervention

Which **populations** benefit from our intervention (e.g. shrimps in different farming systems / countries). In the case of animal organisations, this likely includes both the human and non-human animals.

We used the following Headings:

- **Population** (i.e. Shrimps in leased, semi-intensive ponds in Andhra Pradesh, India)
 - This one is useful to fill in at the end after you've worked through the other headings
 - Try to be as specific as possible
- **Geography** (i.e. Country/State)
- **Scale** (ideally quantifiable, or as a rough %)
- **Tractability** (ideally quantifiable, or as a rough %)
- **Crucial Considerations**
- **Economic factors** (i.e. Semi-intensive Ponds / Leased or Owned Ponds, size of pond)
- **Aligned** (i.e. Resistant to help)
- **Interest** (ideally quantifiable, or as a rough %)

5.3.2. Change Mechanisms

Change mechanisms are the key actions or processes that lead to change. They are like the medication that treats worms in students, or the radio broadcasts that provide health information to the public. For best results, the list of change mechanisms should be specific. For example, instead of saying "aeration devices" to improve dissolved oxygen levels for

aquatic farm animals, it may be more effective to say "2hp paddlewheel aerators: 1 per acre". This level of detail can help identify important nuances and prevent ideas from being discarded due to assumptions about the specifics (e.g. 1 aerator per acre may be a promising change mechanism, while 0.5 per acre may not be as effective).

We now want to select the most promising [Causal Factors](#) identified above, along with our [behaviour change mechanisms](#), and finally put them together and brainstorm some change mechanisms.

We used the following Headings:

- **Change Mechanism** - What is the potential change mechanism?
- **Causal Factor(s)** - Which causal factor(s) does the Change Mechanism effect?
- **Evidence** - What evidence is there that this could trigger change?
- **Other Change Mechanisms** - Are there other change mechanisms that would (or need to) work in tandem with this one

For each of the following, it may only be possible to give an intuitive sense of where this idea ranks, so may be best to use a scale that is ideally quantifiable, or as a rough %

- **Scalability** - How scalable is this as a change mechanism, likely driven by which causal factor it affects
- **Tractability** - How easily are we able to make progress with this change mechanism
- **Cost** - How much would it cost to implement this?

5.3.3. Behaviour Change Strategies

There is a really useful framework we can use to help us with brainstorming our Change Mechanisms - SparkWave's [Conditions for Change](#) tool. Effectively this tool breaks down the conditions for change into 10 steps, which themselves are split into three key stages:

1. Making a **decision** to adopt the new behaviour (*steps 1-3 below*)
2. Performing a number of **actions** that comprise the new behaviour (*steps 4-9 below*)
3. Ensuring the **continuation** of the relevant conditions for success as time passes (*step 10*)

The framework outlines a number of behaviour change strategies that can be implemented at each step (way too many to list here). I'd recommend starting with a framework (such as the one below) and filling it in as you read through the Conditions for Change tool, using the suggested behaviour change tools to brainstorm which tools will be most useful to your specific Populations, Problems and Causal Factors.

[Population X](#) has influence over [Causal Factor Y](#) and we need them to help us change [Problem Z](#). To do this we can explore the following promising **behaviour change strategies**:

1. **Considers** changing the behaviour
 - *Behaviour Change Strategies: Lorem*
2. **Desires** changing the behaviour to engage in the behaviour
 - *Behaviour Change Strategies: Lorem*
3. **Intends** to change the behaviour
 - *Behaviour Change Strategies: Lorem*
4. **Remembers** to perform each action
 - *Behaviour Change Strategies: Lorem*
5. **Believes** attempting each action will help achieve a goal
 - *Behaviour Change Strategies: Lorem*
6. **Chooses** to perform each action
 - *Behaviour Change Strategies: Lorem*
7. **Knows** how to perform each action
 - *Behaviour Change Strategies: Lorem*
8. **Has** needed resources to perform each action
 - *Behaviour Change Strategies: Lorem*
9. **Embodies** skills and traits needed to perform each action
 - *Behaviour Change Strategies: Lorem*
10. **Maintains** attributes required to perform future needed actions
 - *Behaviour Change Strategies: Lorem*

5.3.2. Delivery Mechanisms

5.4. Intervention Design

6SQuID step 4: Identify how to deliver the change mechanism(s).

After identifying the potential change mechanisms, we need to figure out how best to deliver them, which is likely to be target-group and context-specific.

For us, this step essentially operated as a big funnel, where we're able to add any and all intervention ideas to a [huge longlist](#). Then based on an initial shallow prioritisation of ideas, move onto a deeper dive on the [shortlisted ideas](#). Ideas that perform strongly on the shortlist move on to us developing a detailed [programme theory](#) to tease out any more assumptions baked into the idea. Finally, we detail any [crucial considerations](#) that would change our intervention.

5.4.1. Intervention Longlist

What are the possible interventions we could test?

For each of the four core pieces ([population](#), [problems](#), [causal factors](#), [change mechanisms](#)) you should now have many options, with each change mechanism being connected to a specific causal factor, each causal factor being connected to a specific problem, etc.

You can now start [brainstorming](#) ideas. We used a [Weighted Factor Model](#) as part of this process, in order to shortlist the most promising intervention ideas.

We used the following Headings:

- **Intervention** - The name of the intervention
- **Description** - A brief description of what the intervention would entail, largely drawing on the [change mechanisms](#).

For each of the following, try and quantify each (even if it's an intuitive sense). For example, for each we gave them a number on a scale of 0-10. We could then "weight" each score by its importance to give an overall score for the intervention idea, and then we could move any idea that scored above a certain number to the [Shortlist](#).

- **Scalability** -
- **Neglectedness** -
- **Tractability** -
- **Cost-Effectiveness** -

For each of the following, it's especially important to give a useful weighting to each problem. For example, a single intervention may positively impact multiple problems, but you want more impact on more important problems to score higher in your ranking (for example, a Water Quality intervention might impact both Dissolved Oxygen levels and Salinity, but we care more about Dissolved Oxygen, so an intervention that affects Dissolved Oxygen only a little will still likely rank higher than an intervention that affects Salinity a lot).

- **Problem 1** - The extent to which the intervention would positively impact this problem (identified in the problem distribution)
- **Problem 2** - The extent to which the intervention would positively impact this problem (identified in the problem distribution)
- etc.

5.4.2. Intervention Shortlist

Which of our possible interventions are we considering testing?

Any ideas that score highly on your Intervention Longlist should end up on your Intervention Shortlist (this will also likely be somewhat of an ongoing open process - people should feel

free to add new intervention ideas to your longlist, and if they score well, they should be able to do a deeper dive through the shortlist tool).

Any ideas that score highly on the above will be discussed as a team, and then these promising ideas will advance to the next stage. The next stage could be either doing further research on the main [crucial considerations](#), or meeting with [experts](#) to discuss the ideas. Or it could move straight on to developing an [implementation plan](#) to test the viability of the intervention

Effectively, this section mirrors [Charity Entrepreneurship's "Shallow Dive"](#) stage of their research process. We want to explore three key areas of the intervention idea, before finally summarising the shallow dive (spending approximately ~30 mins on each section, so ~2 hours total), an example of Iterative Depth, which we introduced briefly in the [Tools](#) section.

We created a document template using a simple table here, where we break down the 4 key areas, and then do some research to answer the question for each.

1. **Evidence Base** - Look at different versions of the idea and try to steelman the best one. (In some cases, you can rule out an idea at this stage).
 - The **scale** of the issue
 - The *evidence* that it **helps animals**
 - **Implementability**
 - **Case studies** of similar work from other organisations.
2. **Cost-Effectiveness Analysis** - A simple CEA (which we'll go into in more detail in the [Impact Analysis](#) stage), consisting of the following (again, in some cases, you can rule out an idea at this stage):
 - **Number** of animals affected
 - **DALY/Welfare Points** affected by the intervention
 - **Probability** of success
3. **Paths to Failure** - Considering the potential paths to failure for each intervention:
 - **Tractability**
 - **Confrontational** (or whether it would upset the animal movement)
 - **Neglectedness** (or whether an existing org does or will cover this)
 - **Externalities** (or whether it has potential negative risks)
 - **Other**
4. **Summary / Next Steps** - Finally, based on the above, we summarise whether the intervention is promising and why (often for the benefit of discussing the idea with a team of decision makers), as well as the next steps (such as moving on to create a [Programme Theory](#))

5.4.3. Programme Theories

What does the theory of change look like for this intervention?

Programme Theories help to depict the intervention (or programme) diagrammatically. They are very similar to our [Theory of Change diagram](#), except in much finer detail, with a specific focus on the intervention we want to test. In fact, for all intents and purposes, they're the same tool. I just find it useful to separate the idea of an overall "broad" Theory of Change, with the more "in-the-weeds" detailed Programme Theories. Once you have developed a detailed Programme Theory that you have [demonstrated is impactful](#), you'll likely develop your new overall Theory of Change using the Programme Theory you develop here. As before, the benefits of this tool is that it allows us to articulate the short-term Activities and intermediate Outputs that we expect to lead to the long-term Outcomes and our final Impact.

In doing so, we can sense-check our thinking (and identify where we lack empirical evidence to support links in our [causal pathway](#)), discuss our program with a range of stakeholders, and finally test our theories in practice (during [viability testing](#)). As with the [Identifying Assumptions](#) sub-section of the Theory of Change, it can also be useful to label every link in the theory with a number to identify it, as that can be a useful reference when discussing assumptions or Crucial Considerations (next).

We added the Programme Theory for each intervention to its section in the Intervention Shortlist document, so that it was altogether in the same place.

5.4.4. Crucial Considerations

What considerations would warrant a major reassessment of our intervention?

Your Programme Theories should have helped to tease out assumptions - so it's now important to understand the [crucial considerations](#) for your intervention.

Essentially for each of your programme theories you should be able to come up with a list of potential crucial considerations (as with the programme theories, I don't have any special template for this, but it probably makes sense to add your crucial considerations to the relevant section in the Intervention Shortlist, just so everything is in the same place, ready for Viability Testing).

6. Impact Monitoring

A lot of the insights in this section are taken from the book "[The Goldilocks Challenge](#)"

This was a hard section to place within the Roadmap, originally it formed part of the Theory of Change section (and then was moved to just after it, when as a sub-section it became a big enough to justify its own section). However, until you've actually got an intervention to monitor, it didn't make a lot of sense to me to spend resources trying to implement an Impact Monitoring system, especially if it would need to be significantly updated after going through the [Intervention Development](#) step. So I finally decided to place it right in the middle of the 6SQuID process, once you've determined some interventions to test. I think on-balance this is the right-call, but I'd certainly be open to being persuaded that it should instead come after the ToC (provided your high-level ToC is detailed enough, and investing resources in monitoring makes sense for your project at that stage).

We can now use the our [Theory of Change](#) / [Programme Theories](#) to identify key questions for our Impact Analysis:

- **Activities** - Are you working on what you said you'd work on?
- **Outputs** - What tangible outputs did you produce?
- **Outcomes** - Did your outputs achieve their intended short-term result?
- **Impact** - Did your outcomes improve the world?
- **Cost-Effectiveness** - How much good did your impact produce?

Essentially, we can use our Theory of Change diagram to understand what data we should collect in order to assess our impact. As a result, there will be two key outputs that we'll create at this stage, our [Monitoring & Evaluation](#) system, and our [Cost-Effectiveness Analysis](#). In both cases, we'll begin with a simple iteration of each, which we'll then proceed to reflect on after each subsequent stage of the roadmap. Before we develop our two outputs though, I want to quickly discuss [data collection](#) in general..

As a quick aside - The Impact Monitoring section of this Roadmap (much like the [Theory of Change](#)) - **should be referred to (and if needed, updated) often**. It should function as a living document. Ideally, **it should be returned to after every subsequent (major) stage in this roadmap** (i.e. [Viability Testing](#), [Impact Testing](#), [From Growth to Scale](#)). From here on out, your Impact Monitoring system(s) are going to be your trusty compass, and should become more detailed (or streamlined) depending on your intervention(s) needs.

6.1. Data Collection

6.1.1. The CART Principles

The [CART Principles](#) refer to data that is:

- **Credible**: Collect high quality data and analyse the data correctly
- **Actionable**: Commit to act on the data you collect

- **Responsible:** Ensure the benefits of data collection outweigh the costs
- **Transportable:** Collect data that generate knowledge for other programs

The principles apply to any [data collection](#) you may end up carrying out. But they should be used as a guide at any point you're considering data collection, so as to not make the mistake of overinvesting in data, paying for data that is never used or acted on. Essentially you're trying to get the "Goldilocks-ratio" of data - just right 😊

For us, this meant working our way through the following questions about each potential metric, removing any parameters that didn't meet The CART Principles.

- **Credible:** Collect high quality data and analyse the data correctly
 - **Is the data valid?** (does it capture the actual essence of what we're seeking to measure, and not a poor proxy for it?)
 - **Is the data reliable?** (is it possible to gather this data in a way that reduces randomness?)
 - **Is the data unbiased?** (Are we able to reduce bias in the data, particularly in the case of data that can only be gathered through surveys?)
 - **Can we appropriately analyse this data?** (Are we able to measure this data against a counterfactual? I.e. can we get a control)
- **Actionable:** Commit to act on the data you collect
 - **Is there a specific action that we will take based on the findings?**
 - **Do we have the resources necessary to implement that action?**
 - **Do we have the commitment required to take that action?**
- **Responsible:** Ensure the benefits of data collection outweigh the costs
 - **Is there a cheaper or more efficient method of data collection that does not compromise quality?** (i.e. can we take certain data points less frequently and not compromise information value? Can one datapoint stand as a reliable proxy for others?)
 - **Is the total amount of spending on data collection justified, given the information it will provide, when compared to the amount spent on other areas of the organisation?** (e.g. administrative and programmatic costs)
 - **Does the information to be gained justify taking a beneficiary's time to answer?**
 - **Is the added knowledge worth the cost?** (i.e. how much do we already know about the impact of the program from prior studies, and thus how much more do we expect to learn from a new impact evaluation?)
- **Transportable:** Collect data that generate knowledge for other programs
 - **Will future decisions, either by this organisation, by other organisations, or by donors, be influenced by the results of this study?**
 - **Will this data be relevant to the design of new interventions, if required?**

- ***Will this data support the scale-up of interventions that work?***
- ***Does this data help us answer questions about our Programme Theory?***
(allowing other organisations to judge whether the program might work in their context)
- ***Would collecting this data also be useful in other contexts?*** (i.e. are we able to replicate this somewhere else, which if successful, would bolster the relevance of this intervention in multiple contexts)

6.1.2. How to Measure Anything

This section is essentially a summary of the book [“How to Measure Anything”](#).

The concept of [Applied Information Economics \(AIE\)](#), can be summarised as a philosophy in the following six points:

1. If it's really that important, it's something you can define. If it's something you think exists at all, then it's something that you've already observed somehow.
2. If it's something important and something uncertain, then you have a cost of being wrong and a chance of being wrong.
3. You can quantify your current uncertainty with calibrated estimates.
4. You can compute the value of additional information by knowing the “threshold” of the measurement where it begins to make a difference compared to your existing uncertainty.
5. Once you know what it's worth to measure something, you can put the measurement effort in context and decide on the effort it should take.
6. Knowing just a few methods for random sampling, controlled experiments, or even just improving on the judgement of experts can lead to a significant reduction in uncertainty.

In practice, “measuring anything” can be broken down into 5 steps:

- 1. Start by decomposing the problem -**
- 2. Ask is it business critical to measure this? -**
- 3. Estimate what you think you'll measure -**
- 4. Measure just enough, not a lot -**
- 5. Do something with what you've learned -**

6.1.3. Data Collection principles

Chapter 7 of [“The Goldilocks Challenge”](#) was used for the practical nuance in this section.

First-up, a couple of terms it's useful to define. The first one is **Concepts**, which refers to *defining exactly what the organisation wants to measure (i.e. Activities, Outputs, and Outcomes), clearly and precisely*. This will probably require some refinement of your [definitions](#) (i.e. if you're measuring an improvement, what constitutes an improvement to your organisation? Is it 10%? 50%? 100%? If you're measuring a change in behaviour, what constitutes a "change"). The second term we want to define is **Indicators**, which refers to *the metric that will measure the Concept* (we discuss more about how to form your Indicators later in this section).

A first consideration you'll likely want to make is, are we going to use **Primary data** (i.e. data we're collecting ourselves) or **Secondary data** (i.e. data that exists already and we're using), or a mix. There are some obvious trade-offs, with primary data giving you more organisational control over the data being collected, and enabling it to be highly targeted, however it has a high cost (it requires time to develop and test collection methods, train staff, and ensure data quality). Secondary data on the other hand is often much cheaper, but far less targeted, often unavailable in the context you're working in, and the data will still need to be validated for accuracy.

Secondly, you'll need to consider concepts of Validity & Reliability. Validity refers to both **Construct Validity** (*an indicator captures the essence of the concept you are seeking to measure - the output or the outcome*) and **Criterion Validity** (*the extent to which an indicator is a good predictor of an outcome, whether contemporaneous or in the future*), and **Reliability** refers to *data that is consistently generated - in other words, the data collection procedure does not introduce randomness or bias into the process*. Which moves us neatly on to a discussion of Bias and Error...

Bias (in the context of data collection) refers to the systematic difference between how the average respondent answers a question and the true answer to that question. Some common forms of measurement bias are: a) the **Mere measurement effect** (people changing their behaviour because they were asked a question), b) **Anchoring** (a person's answer is influenced by another number or concept), c) **Social desirability bias** (a person's response to a question is influenced by wanting the answer to be the socially desirable answer), and d) the **Experimenter demand effect** (people answer in ways as to please the experimenter).

Error (in the context of data collection) refers to data issues caused by noisy, random, or imprecise measurement of data. Some common forms of measurement error are: a) the **wording of questions** (lack of precision and clarity in questions leading to different responses and unreliable data), b) **recall problems** (the recall period is too long for the context - also known as the "look-back period". You can often consult the literature for appropriate look-back periods in different contexts), and c) **translation issues** (poor or inconsistent translation that fails to account for cultural contexts).

Next we can move on to defining our **Indicators** (taking the “What” from our [data requirements](#), we now want to define our “How”). Indicators should be (which should be easy enough to remember with the snappy acronym *SSTFF*...):

- **Specific** - enough to measure the Concept you're trying (and only that concept).
- **Sensitive** - to the changes it is designed to measure, so that the range of responses is calibrated to the likely range of answers.
- **Time-Bound** - relating to a specific time-period with a defined frequency
- **Feasible** - within the capacity of the organisation to measure (and the respondents to answer)
- **Frequency-calibrated** - dependent on the data that is being captured.

Finally, it's worth considering some **Logistics**, such as: a) **Data collection instruments** (it often makes sense to pretest instruments in both an office and in the field), b) **Survey setup** (ensuring there is a detailed data collection plan, that documents are translated appropriately, data collection staff are trained, and there is good oversight in place - i.e. accompanying, spot-checks and audits), c) **Personnel** (Staff incentives don't unintentionally skew the data, program staff understand the importance of the handling the data and how to do so), and d) **Data Entry** (dependent on paper or digital-based, but embedding concepts like logical checks (not enabling a response that doesn't make sense), and double entry (having two people independently enter paper data electronically and then having any discrepancies flagged electronically)).

6.2. Monitoring & Evaluation

Monitoring & Evaluation (M&E) are two sides of the same coin, to such an extent that it's often easy to forget that they're actually referring to two distinct ideas that work in harmony. As a result, I'm going to divide up my discussion of them here into sub-sections (because who doesn't love division 😊).

It's also worth noting here that M&E systems often don't get it right the first time. Rather they evolve over time, based on internal learning (and external accountability), in addition to evolving along with the given stage of the impact roadmap.

6.2.1. Monitoring

Regular collection of data for real time adjustments to an intervention.

Chapter 5 of [“The Goldilocks Challenge”](#) was used for the practical nuance in this section.

Your monitoring system(s) should a) demonstrate accountability and transparency (that the program is on track and delivering outputs), and b) provide actionable information for learning and improvement (for decision makers).

In practice, this likely means collecting five types of data:

1. **Financial** - Costs (staff, equipment, transport etc.) and revenues (from donors, or a contract, or income if a social enterprise). Often this data is collected anyway as a result of external accountability requirements, however data can also be organised in such a way to implement learning, such as cost-per-output of programs (more in the [Cost-Effectiveness Analysis](#) section).
2. **Activity Tracking** - Key activities (and outputs) identified by the Theory of Change. As with financial data, this is typically collected anyway for external accountability, but often can be made more actionable, by satisfying the following conditions: a) by disaggregating activities across programs/sites, b) by connecting activity data to financial data, and c) regular reviewing of the data to make decisions. High-performing areas can then be learned from and the learnings transferred to others (internal programs, or other organisations).
3. **Targeting** - Tracking if target [Populations](#) are being reached, so that corrective actions can be taken in a timely manner if not. This data likely falls within your Outcomes, or Impact range within your Theory of Change.
4. **Engagement** - The *Extensive Margin* (the size of the program as identified by participation) and the *Intensive Margin* (how intensely did they participate). This is a little trickier when your beneficiaries are non-human animals, but not impossible (and often information around the stakeholders who need to implement your intervention is as impactful, for example, knowing the "% take-up" of farmers approached is useful Extensive margin data, whereas "usage of SWP WQ monitoring services" is useful Intensive margin data). Essentially, once activities are in place and running smoothly - is the program working from a participant perspective?
5. **Feedback** - From participants, which can often shed light on other data, such as Engagement. Again, this is tricky with non-human animal beneficiaries, where proxies may need to be used (such as Operational / Behavioural welfare indicators) as well as feedback from participating stakeholders (i.e. farmer surveys). Feedback data can also include program staff. Feedback can additionally be gathered through [A/B Testing](#) of an ongoing program.

All this data, then, needs to be collated into a system to enable sharing of the data and insights, for accountability purposes (such as with donors, or for the Impact page of your website), or with internal decision makers. Additionally, the systems should facilitate the timely usage of the data. The data shouldn't "disappear into a black hole". Your actionable data system then, should do three things:

1. **Collect the right data** - Following the CART Principles, your data should be Credible (both valid and reliable), Actionable (not collection "easy to collect" data that doesn't have a clear purpose or associated action), Responsible (data use is greater than the costs associated with collection) and Transportable (can produce shareable data and

insights with others, often bringing the costs down, i.e. by using standardised and widely used metrics).

2. **Produce useful data in a timely fashion** - This means being able to resolve differing needs for different insights at different parts of the organisation. The senior team may need aggregated data related to strategic priorities. The Program staff may need disaggregated data to assess program performance. There will likely also be external requirements for the data (such as reporting to donors). The data should therefore be accessible and analysable in multiple ways, and your system should be able to facilitate this.
3. **Create organisational capacity and commitment** - Finally, data should be shared internally, with staff responsible for reviewing and reporting on data, with a culture of learning. Shareable data could be as fancy as a nice digital dashboard, but it also could be as simple as a chalkboard. Similarly, reviewing and reporting on data can be built into regular meeting agendas. Finally, a culture of learning should be built, with staff being able to understand, explain and respond to data - with an emphasis on learning and improvement, not mistakes.

6.2.2. Evaluation

The periodic appraisal of the value of an intervention. Is our intervention relevant, coherent, effective, efficient, impactful and sustainable?

Building on from the [CART Principles](#), we're now moving onto a using (yet another) programme methodology - namely the [OECD DAC Network on Development Evaluation \(EvalNet\) criteria](#) (catchy).

Evaluation is the other side of the "Monitoring & Evaluation" coin, but unlike Monitoring, which is an ongoing process, Evaluation is most useful when scheduled as a periodic appraisal. The idea of this step is to think upfront about how you will evaluate your programme (which should then inform your data collection SOPs, or intervention design). In particular, it's useful to specify ahead of time what actions should be taken based on findings, with an organisational commitment (and resources) to do so. Such as:

- **What circumstances warrant a scale-up of the intervention?**
- **What circumstances warrant a small (or significant) redesign of the program?**
- **What circumstances would warrant a program shutdown?**

It often doesn't make sense to conduct an Evaluation during the [Viability Testing](#) phase (where the overall goal is just to do small-scale tests on different interventions to understand their viability in context). During this point, your routine [Monitoring](#) systems should be enough to allow you to keep tinkering and updating your intervention as you go, until you're confident you have a viable intervention and you want to do [Impact Testing](#). At

which point, it makes sense to do an Evaluation (but for the sake of keeping Monitoring and Evaluation together in the same section of the Roadmap, I'm going to keep this here).

An Evaluation may rely on data you're collecting during Business as Usual [Monitoring](#), or it may warrant a larger data collection effort with a new SOP created to capture data required only for the Evaluation, such as when conducting a Randomised Control Trial or Quasi-Experimental Methods. A good starting point on determining whether an Evaluation makes sense at a certain point is to again turn back to the CART Criteria:

- ☐ **Are we able to credibly estimate the counterfactual?** (Would the evaluation be *Credible*?)
- ☐ **Are we committed to using the results, whether they are positive or negative, and do we have the resources to do so? Is the timing right?** (Would the evaluation be *Actionable*?)
- ☐ **Do the benefits of impact evaluation outweigh the costs? Would this evaluation add something new to the available evidence?** (Would the evaluation be *Responsible*?)
- ☐ **Would the evaluation produce useful evidence for others?** (Would the evaluation be *Transportable*?)

However, while the CART Principles on their own would likely suffice for developing an Evaluation Process, we found that the [OECD EvalNet criteria](#) contained a plethora of specific questions that could be asked in order to further refine your Evaluation. The EvalNet methodology uses six criteria, with two principles for their use.

The two principles are:

1. **The criteria should be applied thoughtfully to support high-quality, useful evaluation** - *They should be contextualised – understood in the context of the individual evaluation, the intervention being evaluated, and the stakeholders involved. The evaluation questions (what you are trying to find out) and what you intend to do with the answers, should inform how the criteria are specifically interpreted and analysed.*
2. **The use of the criteria depends on the purpose of the evaluation. The criteria should not be applied mechanistically** - *Instead, they should be covered according to the needs of the relevant stakeholders and the context of the evaluation. More or less time and resources may be devoted to the evaluative analysis for each criterion depending on the evaluation purpose. Data availability, resource constraints, timing, and methodological considerations may also influence how (and whether) a particular criterion is covered.*

The six evaluation criteria are:

1. **Relevance** - *Is the intervention doing the right things?* (The extent to which the intervention objectives and design respond to beneficiaries , global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change)
2. **Coherence** - *How well does the intervention fit?* (The compatibility of the intervention with other interventions in a country, sector or institution)
3. **Effectiveness** - *Is the intervention achieving its objectives?* (The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups)
4. **Efficiency** - *How well are resources being used?* (The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way)
5. **Impact** - *What difference does the intervention make?* (The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects)
6. **Sustainability** - *Will the benefits last?* (The extent to which the net benefits of the intervention continue, or are likely to continue)

In practice, using these criteria looks like asking (and answering) a number of questions about the intervention. We can create a document with four different sections - 1. Principle One questions, 2. Principle Two questions, 3. Choosing which of the Six Criteria to use, and 4. The Evaluation Criteria questions.

Below are some example questions (all taken from the OECD book, [available online](#)):

Principle One (the criteria should be applied thoughtfully to support high-quality, useful evaluation):

- **Context:** What is the context of the intervention itself and how can the criteria be understood in the context of the individual evaluation, the intervention and the stakeholders?
- **Purpose:** What is the evaluation trying to achieve and what questions are most useful in pursuing and fulfilling this purpose?
- **Roles and power dynamics:** Who are the stakeholders, what are their respective needs and interests? What are the power dynamics between them? Who needs to be involved in deciding which criteria to apply and how to understand them in the local context? This could include questions about ownership and who decides what is evaluated and prioritised.
- **Intervention:** What type of intervention is being evaluated (a project, policy, strategy, sector)? What is its scope and nature? How direct or indirect are its expected results? What complex "systems thinking" are at play?

- **Evaluability:** Are there any constraints in terms of access, resources and data (including disaggregated data) impacting the evaluation, and how does this affect the criteria?
- **Timing:** At which stage of the intervention's lifecycle will the evaluation be conducted? Has the context in which the intervention is operating changed over time and if so, how? Should these changes be considered during the evaluation? The timing will influence the use of the criteria as well as the source of evidence.

Principle Two (the criteria should not be applied mechanistically):

- What is the demand for an evaluation, who is the target audience and how will they use the findings?
- What is feasible given the characteristics and context of the intervention?
- What degree of certainty is needed when answering the key questions?
- When is the information needed?
- What is already known about the intervention and its results? Who has this knowledge and how are they using it?

Choosing which of the **Six Criteria** to use:

- If we could only ask one question about this intervention, what would it be?
- Which questions are best addressed through an evaluation and which might be addressed through other means (such as a research project, evidence synthesis, monitoring exercise or facilitated learning process)?
- Are the available data sufficient to provide a satisfying answer to this question? If not, will better or more data be available later?
- Who has provided input to the list of questions? Are there any important perspectives missing?
- Do we have sufficient time and resources to adequately address all of the criteria of interest, or will focusing the analysis on just some of the criteria provide more valuable information?

Finally, you'll be at a point where you can attempt to answer (or get a sense of how you'll answer) the actual **Evaluation Criteria** relevant to your intervention. Remember that you needn't, or even shouldn't use ALL of the criteria. Questions you might be answering at this point are (as above, these are all taken directly from the OECD, and in this case, all use a gender lens, which I thought would provide a nice break from using shrimps as the example every time):

1. **Relevance** - *Is the intervention doing the right things?*
 - Was the intervention designed in ways that respond to the needs and priorities of all genders? If so, how?

- To what extent does the intervention's design reflect the rights of persons of all genders and include feedback from a
- diverse range of local stakeholders including marginalised groups?
- Does the intervention meet the practical and strategic needs of all genders?
- 2. **Coherence** - *How well does the intervention fit?*
 - To what extent are the intervention's design, delivery and results coherent with international laws and commitments to gender equality and rights, including the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW), the Beijing Declaration and Platform for Action, the Programme of Action of the International Conference on Population and Development, and the 2030 Agenda?
 - To what extent does the intervention support national legislation and initiatives that aim to improve gender equality and human rights? What lessons can be learned?
- 3. **Effectiveness** - *Is the intervention achieving its objectives?*
 - Did the intervention achieve its objectives and expected results in ways that contribute to gender equality? If so, how?
 - Were there differential results for different people? If so, how and why? Were different approaches necessary to reach people of different genders? Was there sufficient monitoring and analysis of differential effects? Was the intervention adjusted to address any concerns and maximise effectiveness?
 - Was the theory of change and results framework informed by analysis of gender equality, political economy analysis and human rights? If so, to what extent?
 - To what extent and why is effectiveness different for people of different genders?
- 4. **Efficiency** - *How well are resources being used?*
 - Were different resources allocated in ways that considered gender equality? If so, how were they allocated? Was differential resource allocation appropriate?
 - Do the investment costs per person targeted meet the differentiated needs of people of different genders?
- 5. **Impact** - *What difference does the intervention make?*
 - Were there equal impacts for different genders or were there any gender-related differences in engagement, experience and impacts? If so, why did these differential impacts occur?
 - To what extent did gender-related impacts intersect with other social barriers including race/ethnicity, disability, age and sexual orientation to contribute to differential experiences and outcomes?
 - How did gendered norms and barriers within the wider political, economic, religious, legislative and socio-cultural environment impact outcomes?

- To what extent have impacts contributed to equal power relations between people of different genders and to changing of social norms and systems?
- 6. **Sustainability** – *Will the benefits last?*
 - Did the intervention contribute to greater gender equality within wider legal, political, economic and social systems? If so, how and to what extent? Did it result in enduring changes to social norms that are harmful to people of all or some genders? If it did not achieve this, why not?
 - Will the achievements in gender equality persist after the conclusion of the intervention? Have processes contributed to sustaining these benefits? Have mechanisms been set up to support the achievement of gender equality in the longer term?

Phew! That was a lot of questions! But it should leave you in a much better position to understand where your key uncertainties lay - and what kind of data you'll need to start collecting in order to answer these questions.

6.2.3. Logframe

A good place to start with M&E (well, [Monitoring](#) in particular) is a logframe. A logframe essentially takes your [Theory of Change](#) (or a [Programme Theory](#)) and transforms it into a table, allowing you to inspect each step, and pull out the data that can be used to verify that the change is taking place. You take each step (Objective) and then explicitly state what you're going to track (Indicators/Metrics), how you'll verify and store the data related to this tracking (Sources of Verification), and finally any risks, assumptions or externalities within this step, that may result in no impact, even if the Objective is achieved (Threats). While your final Monitoring & Evaluation system may end up outgrowing the Logframe, this is a great place to start.

With the Sources of verification in particular, you can specifically link to where the data will be stored. Ideally in the future, all these data sources could ultimately pull into one single "dashboard" of data, but to start with, just knowing where the data will be accessible is enough.

Your logframe will look something like this (I've filled in an example of what this will look like, focusing on the Supply-side work of SWPs (overall) Theory of Change - *a fully completed Logframe would include all steps in your ToC*):

	Objective	Indicators	Verification	Threats
Activities	Reach out to shrimp producers	# of calls with shrimp producers	Meeting notes stored in SWPs Notion (filtered by	- Meetings, calls and gatherings not translating into real

		# of producers we speak to / meet	"Producers") here	interest in participating
Outputs	Work with producers to commit to our Asks (i.e. Water Quality, Stocking Density limits)	- # of agreement(s) reached to commit to our Asks	Signed documents with producers (stored in Google Drive folder here)	- Failure to comply - Corruption - Farmers disinterested
		- # of Shrimp affected by commitment		
Outcomes	Our Asks are implemented and monitored on farms	Water Quality metrics	Real data captured with probes and stored in databases here	- Manipulation of data - Unable to access data
		Other M&E metrics if appropriate (i.e. stocking density or ablation)		
Impact	Shrimp are better able to breath	Dissolved oxygen concentration	- Data captured from farms stored in databases here (AI Apps looks potentially useful for lots of metrics) - Eventually RCTs	- Increased stocking densities (*) - Increased demand for shrimp - Research conclusively proves shrimp aren't sentient / can't suffer - Body of knowledge is less extensive than for other animals
	Reduced mortality rates	Mortality rates		
	Lower incidence of disease	Disease rates		
	Shrimp are less stressed / anxious	Growth rates		
		Feed conversion ratios		
	Improved environmental conditions	Stocking density		
	Increased human preference from behind the veil of ignorance	(above, plus) Slaughtering method utilised		
		Eye-ablation rates		
Cost - Effectiveness	Improved welfare of whiteleg shrimp cost-effectively	Total number of shrimps helped	- Data captured from farms (offset by Welfare Points) here	Counterfactually strong funding
		Total number of	- Data captured from	Counterfactuals of

		shrimps saved	farms (offset by Supply & Demand analysis) here	co-founders
		Shrimps Impacted per \$	- Actual costs incurred and impact measured here - Cost-Effectiveness Analysis	

6.3. Cost-Effectiveness Analysis

More details on developing a Cost-Effectiveness Analysis can be found in chapter 14 of the Charity Entrepreneurship Handbook - [“How to Launch a High-Impact Nonprofit”](#) (and also likely elsewhere...)

A Cost-Effectiveness Analysis (CEA) is basically a calculation aimed at creating a single number representing a benefit (such as a health metric, like DALYs or Lives Saved) for a given cost (normally some currency). **E.g. “the result of the CEA showed a life could be saved for “\$5,000”.**

The CEA, as with the rest of these Impact Monitoring tools in this section, is likely to be a tool that you create a simple, first iteration of, and then routinely review and update (either at set intervals, i.e. after x months, or at predetermined milestones, i.e. after x has finished). Therefore, it's a little hard to write a guide for the level of detail you'll need, it's likely you'll design a first iteration of your CEA with less detail than is described below, and you'll likely end up with a 'final' iteration with more detail than is described. Additionally, it's worth remembering that the CEA is [a map, not the territory](#). It's likely to be imperfect and/or missing key data. Whilst it's often a very useful tool, there are shortcomings, and it should be one tool among a few to help you make decisions (*on this note once you've finished your CEA, it's worth referring to this [list of ways in which cost-effectiveness estimates can be misleading](#), it's a useful checklist to refer to in order to assess if important considerations haven't been missed*).

Additionally, bear in mind that there is diversity among existing CEAs. Charity Entrepreneurship suggests that there are the following levels of CEAs:

- **Informative** - A CEA where the endline isn't particularly useful on its own, but the variables and citations that were used to develop it can be informative for your own.
- **Suggestive** - A good-quality CEA, but assessing a different metric/situation to you. Often it will help you to update your views, but you'll still need to create your own CEA.

- **Predictive** - A high-quality CEA that is sufficiently close to your metrics/situation. You'd likely use many of the same inputs when creating your own CEA.

With those caveats in mind, let's dive into creating your own CEA 😊

You'll likely want to start with a spreadsheet tool like [Google Sheets](#) (a tool like [Guesstimate](#) can be used for complex CEAs you may develop in future, in particular Guesstimate is useful for CEAs with high levels of uncertainty).

The design of your CEA will likely split up each **broad idea** (i.e. "number of shrimps farmed per year" or "[Welfare Points](#) associated with maintaining Dissolved Oxygen levels") within your CEA into its own sheet. Each sheet will likely be read across a given row, with the first column containing your titles. You might put specific sections into their own boxes (for readability) and notes in the last columns. Once you've developed your sheet for each broad idea, the key number from each can be pulled into a **Summary Sheet**. The summary sheet will contain the results of the CEA, i.e. "Benefit X for Cost Y ", as well as the most important numbers that factor into that result and the factors which most affect the estimate (the **Sensitivity Analysis**). You'll likely include a **Benefits, Costs and Counterfactuals** section. Additionally, you might also want to include an **Optimistic, Pessimistic, and Best-Guess** for some numbers (for Monte Carlo simulations you can use Guesstimate). Finally, you'll want to include a **References** tab at the end (along with links to the most important numbers actually within the CEA itself).

You might also discount numbers within your CEA, such as **Certainty** (if you have a source for a number, but it's uncertain, you might apply a certainty discount to it), **Generalisability** (when generalising evidence from one context to another, you might apply a generalisability discount, to acknowledge that the situation you're modelling is not identical to the source you're using) or **Bias** (if you believe that bias may factor into a number from a source, you might discount it).

To keep your CEAs consistent with others (like GiveWell's) you'll likely want to use similar formatting, such as colour coding cells based on data type:

- **Yellow: Value and Ethical judgements** - These are numbers that could change if the reader has different ethical judgements to you, and there is often no clear answer - i.e. "how many years of happiness is losing the life of one child under five worth"
- **Green: Citation-based numbers** - Numbers based on a specific citation, often hyperlinked (or noted in the reference section)
- **Blue: Calculated numbers** - Numbers generated from other numbers within the sheet
- **Orange: Estimated numbers** - Numbers where a source cannot be found, and a number is required to be estimated by a CEA author (or expert)

7. Viability Testing

6SQuID step 5: Tests and refinements on a small scale.

In a way, designing an intervention is like a funnel, comprising three overall phases of work:

1. First, you're trying to generate a list of all possible interventions that could be theoretically impactful, and whittling these down to the ones which are the most promising (something like 10+ ideas, ideally lots more) - [Intervention Development](#)
2. Secondly, you're trying to test the most promising of these theoretical interventions in the real-world, to get an understanding of the actual viability of these interventions in the context you're working in (something like 3-5+ ideas) - [Viability Testing](#)
3. Thirdly, the interventions that have now demonstrated themselves to both be theoretically promising and viable in your context, now need to be tested to assess their actual impact, with the aim to then scale-up the most cost-effective intervention (something like 1-3 ideas) - [Impact Testing](#)

Essentially everything up until this point has been part of the “**theoretical**” phase of work. The next stages constitute the “**viability**” phase, before finally we can move on to the “**impact**” phase.

During this Viability phase, the feasibility of our potential interventions need to be tested (and adapted). This is sometimes referred to as ‘[Formative Evaluation](#)’ (I liked this term as it was a way to refer to a real-world test without using the word ‘[Pilot](#)’, which comes as part of the Impact Testing phase, and it was important to differentiate the two. Though sometimes informally, we would refer to these as our ‘Pre-Pilots’).

It may turn out that some of our interventions are not viable (in this context) in which case, they should be removed at this stage. As a result, we should **prioritise tests that can shed light on multiple major crucial considerations**. Additionally, we should aim to anticipate possible unintended effects of the intervention and minimise any that might be harmful.

[Monitoring](#) the viability of the intervention on a small scale can provide important information/learning that may have otherwise been difficult to gain, such as:

- Whether it is accepted by the target group, practitioners, and organisations responsible for implementation
- The most effective elements such as level of participant involvement, format, and length
- Who should be responsible for delivering the intervention and in which settings
- The training required for those involved in delivery
- How to maximise [Population](#) reach

I think it's important to re-emphasise the difference I mean between [Viability Testing](#) and [Impact Testing](#). With viability testing, you're drawing on all the knowledge you've acquired up to this point, in order to test whether your intervention(s) work in the context you're trying to work in, and learn as much as you can as you do it. In Impact Testing, you're scaling up tests on interventions you're confident in, in order to assess their effectiveness in order to justify full-scale implementation (or investment in a large-scale evaluation, such as an RCT). Essentially, it's important to think of the Viability Testing stage as the "[Lean Startup](#)" phase of your intervention - the goal is to test your hypothesis, iterate, and achieve validated learning (with the strong caveat that you should have done the steps up until now and so have a strong theoretical understanding of the problem you're trying to solve, as well as thoughtfully engaged with the ways that the intervention could cause potential harm. There's a lot we can learn from the for-profit world, but we need to ensure that when applying these principles to the non-profit sector, we do so thoughtfully).

7.1. Implementation Plan(s) (Formative Evaluation)

What is our plan for the interventions we want to test?

The implementation plan requires clarifying the conditions and resources necessary for successful implementation and the related risks and assumptions.

A key difference between the Implementation Plan you'll create here (for the [Formative Evaluation](#), or "Pre-Pilot") and the Plan you'll create in the next phase (for the [Pilot](#)), is that the primary aim of the Formative Evaluation is the learning value (does this intervention work in this context? Are there any tweaks we can make along the way to improve it?). Whereas the primary aim of the Pilot is to assess the (cost)-effectiveness of our intervention(s). The Implementation Plans in both cases will likely look somewhat similar, but with this one, you'll likely want to build more slack into the plans, to allow for things to go wrong, or to change the way you're doing things along the way. You'll also likely spend less on data collection, or prioritise a few proxy data points during this phase, as you're essentially just trying to figure out if this intervention is **Viable** in this context (and promisingly impactful), rather than doing a more focused look at evidence of effectiveness, which will come next (though obviously if it's clear the intervention isn't impactful at this point, even if it is viable, you should shut it down).

For us, this was a Google Doc, with the following sections:

- **Project Scope** - A sentence (or short paragraph) containing the Project Objectives, Deliverables and Acceptance Criteria.
- **Programme Theory** - This can be taken directly from the [Programme Theory](#) you created in the Intervention Design stage. This also includes the same process of node labelling for [Identifying Assumptions](#) as with your Theory of Change

- **#** - Which step(s) in the Programme Theory does this relate to?
- **Note / Risk / Assumption / Externality** - Is this a general **Note** about the step? Or a **Risk** (something that could go wrong, likely pulling from our Crucial Consideration)? Or an **Assumption** (an underlying assumption that we want to surface)? Or an **Externality** (or a potential positive or negative result that may occur)
- **Description** - What is the Risk / Assumption / Externality we have identified
- **Likelihood** - What is our expected likelihood of this happening (we used a %)
- **Impact** - What is the potential impact if this happens (we used Small / Medium / Large)
- **Mitigating Actions** - What mitigating actions can we take to prevent this from happening (or what will we do if this happens)
- **Deliverables and Schedule** - A way to track the overall deliverables of the project, as well as a timeline for the results. These could vary significantly depending on your intervention, but I want to provide examples if I can, which can be found in the following section: [\(Potential\) Deliverables](#).
 - **Key Result** - What is the key result we want to see
 - **Deliverable** - What is the deliverable that will help us achieve this key result (multiple deliverables can be assigned under a single key result)
 - **Timeline** - When do we expect this deliverable to be finished. We used rough dates, such as a month, rather than a specific date, but this can be tailored to your intervention. We also used a rough colour coded grading to track the progress of this at-a-glance (Blue meant Completed. Green meant On Track. Amber meant Falling Behind. Red meant Off Track).
 - **RACI** - Who is **Responsible** (for delivering i.e. a data collector), **Accountable** (for successfully addressing it -i.e. a manager), **Consulted** (for advice or input) and **Informed** (of progress).
- **Resource Allocation** - The allocation of resources (time and money) dedicated to the project
 - **Owner** - Who is responsible for this resource
 - **Requirement** - What is the resource requirement (i.e. time spent doing X, or purchasing of equipment Y)
 - **Cost** - Likely either time spent (in hours) or cost (in dollars)
 - **Project Costs** - Based on the data above, you should be able to calculate an overall cost for the project (including costs associated with time spent, such as that person's hourly pay), feeding into the Responsible [CART principle](#)
- **Appendix** - Anything else. We found it useful to include all the SOPs at the end, so everything was available in the same document (even though in future, it makes sense to have the SOPs as separate, shareable documents)

7.2. (Potential) Deliverables

The following is a list of potential deliverables that could be part of your [implementation plan](#). It's largely based on the implementation plan we created, and yours might look quite different (depending on the intervention(s) your testing). However the overarching point should be clear, that the implementation plan isn't the only document you'll be creating at this stage, but rather, the implementation plan will be your guide in helping you determine the deliverables that will be required during your testing phase.

7.2.1. Intervention Delivery SOP

The Standard Operating Procedure (SOP) outlining the delivery of the intervention (potentially a one-off, or ongoing). For us, this looked different depending on the intervention delivery being tested.:

- **Corporate Campaigns** - We wanted to test the tractability of corporate campaigns for shrimp welfare. This is something we were initially weary of, but some initial [expert feedback](#) gave us hope that this could be promising. We decided to test this using the framework of [Outcome Mapping](#). If Outcome Mapping sounds like a useful tool for your intervention, I would recommend the [full book](#).
- **Lowering Stocking Densities** - We wanted to test to what extent lowering stocking densities improves outcomes (such as better water quality, larger shrimps etc), so in this case our "intervention delivery" was in fact a "one-off". We stocked a pond with a lower density than a control pond. In this case, the rest of the deliverable was in ensuring a robust [Data Collection](#) protocol
- **Online welfare oriented advice** - Our [Needs Assessment](#) highlighted that there were a large number of Indian Shrimp Farmers who used social media to get advice on shrimp farming. We wanted to test if we could use this as a method of outreach and information dispersal, offering welfare-oriented advice to farmers. For this, we needed to develop a plan for the ongoing delivery of this intervention (as well as the initial scope)

7.2.2. Data Requirements

What data do we need to gather in order to test our intervention?

At this point you're just trying to get an understanding of what data you need - the "**What**" of the data (i.e. to get the depth and breadth "just right" in order to make a decision or understand a crucial consideration). It's during the next step ([Data Collection SOP](#)) where you'll need to actually formulate the plan for how to gather this data - the "**Who, When, and How**" (the "**Why**" of the [data collection](#) was answered in the "[The CART Principles](#)" section). This stage could take many forms depending on the intervention(s) you want to test, so I'm

hoping the structure I provide here is useful and generalisable, though your intervention(s) may look very different to ours.

We used a longlist/shortlist process, similar to the process followed in [Intervention Design](#).

Longlist: We started with a longlist of all the potential pieces of data we could collect, then prioritised them into a shortlist of data we thought would be the most useful. We used the following Headings:

- **Domain** - As in previous examples, sometimes it's useful to group together similar data points, such as Water Quality Data, Survey Questions)
- **Parameter** - The actual parameter/data point we could collect (i.e. Dissolved Oxygen, or Pond Size).
- **Importance** - Largely pulled from the [Evidence Review](#), how important is this parameter to collect (in some cases, this might be a direct 1:1, so Dissolved Oxygen is important for welfare and is something we can measure, so this is clearly High importance. In other cases it might not be a direct ratio, for example, we might want to collect information about the Weather or Time of Day when a reading is taken, in order to see if they correlate with water quality fluctuations).
 - **Fluctuate** - This was something specific to our longlist for Water Quality as we knew that some parameters fluctuate significantly, some fluctuated somewhat, and others only fluctuated over days or weeks. We wanted to record this in order to help inform the Frequency and Testing Method.
 - **Frequency** - How often we want the data (i.e. something like "Continuously", "Twice Daily", "Weekly", "Monthly", and "Just Once"). If your parameters don't have different frequencies associated (for example, you're just doing a one-off survey or something), you likely don't need this heading.
- **Testing Method(s)** - What are the potential ways to get this data (i.e. there are multiple ways to get the same data for water quality, some are more expensive and give more frequent/accurate data, some are cheaper but provide rough approximations. You can narrow down which testing methods you're going to use in your [Data Collection SOP](#) next - right now it probably makes sense to list them all so you have your options).
 - **Recording Method** - In our case, some data was recorded digitally and some was written down, so we needed to be able to differentiate between the two. If you don't have this consideration, you likely don't need this bit.
- **Reasoning / Notes** - Anything you want to note that doesn't fit with the other sections

Shortlist: You should now create a shortlist of this data using [the CART Principles](#) (This stage should remain focused on the "**what**" of your data requirements. The "**how**" will come next in your [data collection SOP](#)).

At the end of this process, you likely need to get some feedback on your shortlisted dataset, even if just to ensure that the final shortlisted dataset still makes sense as a cohesive whole. Now we can move on to how we're actually going to collect this data.

7.2.3. Data Collection SOP

The theory from the [Data Collection](#) sub-section of [Impact Monitoring](#) should be referred to here.

You likely want to put together some SOPs for your data collection. For example, we had multiple pieces of information to gather, and decided to collect this through the use of multiple digital forms. Some were filled in once only (i.e. when onboarding a farmer), some data was to be collected at twice-daily intervals (such as water quality data that could be gathered with monitors we purchased, and others at weekly intervals (such as water quality data that needed to be analysed at a lab). Your SOPs should make clear the answer to the following questions:

- The breakdown of data should be for each form
- The number of people required for each aspect of each form
- The equipment required for each form
- The frequency of filling in each form
- The expected time to complete each form
- Supplementary information that's useful for the data collector to have (i.e. detailed, clear instructions on how to take a water quality reading)

Additionally, this step will likely include building the form (i.e. unless this is to be collected on a blank notebook you'll likely be designing a custom form and then printing it out, or through use of a digital collection tool like [Jotform](#))

Having a good understanding of your form / data breakdown should help with the next potential deliverable - [Purchasing equipment / Hiring staff](#).

7.2.4. Purchasing equipment / Hiring Staff

Purchasing equipment - An output could just be the purchase of necessary equipment, in which case you can likely just link to the budget or receipt etc. (or a document outlining your procurement process, such as why you decided to buy a specific item).

Hitting staff - Additionally, you may need to plan for an increase in staff in order to carry out your intervention, and this should be planned in (both in terms of time spent recruiting and onboarding, and staff salary).

As with [data storage and analysis](#) below, these two outputs may not require a document of their own, but they should still be factored in as a Deliverable within your Implementation Plan, along with the respective [RACI](#) responsibility assignment.

7.2.5. Data Storage and Analysis

These two should probably be in two separate deliverables, but for the sake of brevity here I'm combining them.

Data storage - You should have a clear outline of how the data will be stored (if you're collecting data electronically, this may be as simple as making sure the data is compiled and backed up to the cloud. But if it's on paper, who is responsible for collecting the data together and processing it, i.e. likely translating it into a spreadsheet - again, get a clear [RACI](#) in place).

Data analysis - Who is responsible for analysing the data? What are the expectations (you're probably trying to answer specific questions in the data as outlined during your CART analysis, rather than a general "exploration" of the data). By when?

7.2.6. Report

As with the [Literature Review](#) and [Context Analysis](#) pieces of work, it probably makes sense to write up your learnings from your viability tests in a report (even if these ultimately remain internal). This both helps to formalise the results (and ensure internal cohesion on what the results mean), but also so that the results can be shared with interested parties (the [Transportable principle](#)) in order to inform other NGOs. It probably makes sense to outline the scope of this upfront (i.e. [Time-Capping](#), is this a 2-page or a 20-page report, or rather, should someone spend 2-hours or 20-hours on it).

7.3. Formative Evaluation(s)

Small-scale tests on our promising interventions.

Now it's time to carry out our Formative Evaluation(s)! The feasibility of our interventions can now be tested (and adapted). I want to emphasise that you're going to learn a tremendous amount at this point simply by doing. It can be easy to fool yourself (given the enormous amount of work you've done up to this point) that your intervention is going to run smoothly, but no plan survives contact with reality.

The point of this step isn't to unwaveringly stick to the implementation plan, it's highly likely that small tweaks can be made to your plan along the way, in which case, go ahead and make those tweaks. The point of this step is to **learn and iterate** - we're going to be doing a robust [Pilot](#) next as part of the [Impact Testing](#) phase. This is not your pilot - so don't be afraid

to make changes to your intervention at this stage - in fact, that's the whole point of this stage!

It may also simply turn out that some of our interventions are not viable in the context you're working in (e.g. something that is highlighted in the literature doesn't work in your context) in which case, these ideas can be removed at this stage.

Considering this is likely to be one of the longest steps in the whole Roadmap, in a way it's surprising how little I have to say about it. But in a way **every step up until this point has been to unlock the knowledge value contained within this Formative Evaluation step**. But it's hard to write the roadmap for this bit because by its very nature, the learning you'll do at this point is currently unknown to us.

It's also worth noting that the relatively short length of this section (compared to the rest of the post) should not denote its relative importance. There's a lot of guidance that can be provided on the theoretical side of things, but in a way you're somewhat on your own at this point. You're going to spend a lot of time doing and learning - don't be surprised if this stage takes as long (or longer) than all the steps up to now put together. This phase needs to be as flexible as your intervention needs until you're confident enough in your intervention that you want to test its impact.

Best of luck - and I'll see you at the Impact Testing stage! 😊

8. Impact Testing

6SQuID step 6: Collect sufficient evidence of effectiveness to justify rigorous evaluation / implementation.

We made it! We now know which of our potential interventions are viable in our context. Before committing to a scale-up (or a more rigorous [evaluation](#), such as an [RCT](#)), we need to establish evidence of effectiveness to warrant the investment. Essentially we're assessing the [\(cost\)-effectiveness](#) of our intervention(s) to find the best one to scale.

8.1. Implementation Plan(s) (Pilot)

Depending on the nature of the results of your [Formative Evaluation\(s\)](#), this step could look very different. For example, if your Formative Evaluation was testing an intervention that:

1. **Is still ongoing** - and seems promising with a few tweaks, then it might just be the case of reusing your original [Implementation Plan](#), updating it somewhat based on the learnings (such as tweaking the [Programme Theory](#), or increasing the number of

beneficiaries, or increasing the scope for [data collection SOP](#) (though still within [the CART Principles](#)) and then having the Formative Evaluation naturally **evolve** into your larger Pilot test.

2. **Was finite** - (as in the case with us, we wanted to test the effects over a single shrimp crop, so when the crop was finished, we finished our [Formative Evaluation](#)), but promising, then similar to the above, it may be the case that you use your original [Implementation Plan](#) as a basis for a new one for the Pilot (likely with some tweaks to the [Programme Theory](#), scope of the test, and [data collection SOP](#) etc). In this case your Pilot will be a **new** project to your Formative Evaluation, but still follow a similar model.
3. **Needs to be tweaked significantly** - but in such a way that you're confident that the new intervention is viable, and you're ready to move onto Impact Testing. For example, you may have been testing a couple of different intervention ideas during your [Formative Evaluation](#), and on their own they weren't particularly viable, but based on some late-stage learnings in the process, you have strong reason to believe that a new intervention that combines the two (or aspects of the two) will be impactful. Or you tried the intervention with a certain [Change Mechanism](#) that wasn't that effective, but the intervention itself seemed to be effective (for example, you may simply pay beneficiaries to behave in a certain way to test whether the intervention worked, before isolating and testing change mechanisms). Either way, it probably makes sense to create a **new** Implementation Plan pretty much from scratch here, with a fresh [Programme Theory](#) and SOPs etc.

All that's to say, that the template for the Implementation Plan(s) here, is going to look quite similar to the one for the Formative Evaluation, so I won't reiterate the template we used, you can just refer back to the [Implementation Plan](#) and [Potential Deliverables](#) of the previous sections. However it's important to think of this as a **new** document, even if you can likely reuse a lot of the original plan. The goal of the [Pilot](#) is different. We now know that the intervention we are testing is viable in the context we're working in. The goal here is to understand the **Impact** of the intervention, in particular the [\(cost\)-effectiveness](#) of the intervention. So there should be a stronger focus on a robust plan which likely will not need to be tweaked/changed (much) once the Pilot has begun. The SOPs you create for [carrying out the intervention](#) and for [data collection](#) should be more robust (you might also explore different data collection methods at this point with a focus on counterfactuals, such as before/after survey of beneficiaries, or implementing a more robust control). You might also include an SOP for scaling down the pilot if it is determined that it is less effective, and you don't want to cause unnecessary harm by not smoothly transitioning away from your active implementation.

It's quite possible you'll just be doing one Pilot here, with an intervention you're pretty confident in - but it's also likely you have a few interventions that seem promising, in which

case you'll be doing 2-3 pilots (potentially at the same time). In fact, that's even potentially preferred, as the goal at this point is to figure out what the most impactful intervention is, in order to scale up.

Essentially the big difference here is that **you're expecting that the intervention you're Piloting here is your organisation's intervention (or one of them)**. So you want to replicate as close to what this intervention will look like as your future Business-As-Usual as possible (of course, you can, and should, still update your intervention based on new evidence during and even after the Pilot, but the intention is to have already unearthed all the learnings you can, so you shouldn't have to). You're likely testing your intervention on a much bigger scale, and the Implementation Plan you create should reflect that. You're also likely trying to calculate an explicit [cost-effectiveness](#) of your intervention that captures nuance, so your implementation plan should also reflect that.

8.2. Pilot(s)

Larger tests on viable interventions.

Congratulations again! You've now got an intervention(s) you're pretty confident is impactful, as well as a plan to test the impact. Now it's a case of following your plan and seeing the results 😊

Similar to the [Formative Evaluation](#) section, considering how important this step is, it feels somewhat strange reading back over this and realising that I actually don't have a whole lot I can write here. Providing your [Implementation Plan\(s\)](#) are robust, at this point it's essentially a case of following the plan.

As a (final) reminder - the goal now is to understand the [cost-effectiveness](#) of your Pilot(s) in order to justify [rolling out your intervention](#) or a more rigorous [evaluation](#) (like an [RCT](#)). Depending on the results of your Pilot(s), you may either decide to move on to [scaling the intervention](#) (if successful), moving back to the [Viability Testing](#) phase (if there is still more learning/tweaking to be done), or even shutting down the project altogether. If you're testing multiple interventions, you may simply find one is the most cost-effective and then scale that one up (in which case, you're hopefully following a "smooth shutdown SOP" for the interventions that will no longer continue).

9. Growth

You've completed your "Explore" phase! Congratulations 😊. Now it's time for the "Exploit" phase (I'm not a huge fan of this terminology, but [Explore/Exploit](#) does succinctly capture a useful idea).

You've determined which intervention is the most impactful, it's time to roll it out. In the beginning of your Exploit phase, you'll likely be in a stage of slow, sustainable growth, finding beneficiaries and rolling out your intervention, increasing your organisational capacity (i.e. staff / equipment / fundraising) along with your growth. Depending on your intervention, a slow and steady [growth](#) might be enough. However, in some cases, you may want to [scale](#) (likely after a somewhat stable period of growth, but I suppose there are situations where you might go straight to scale).

In this section, we're going to come up with a plan to manage your growth, and then can move on (in the next section) to determine a strategy for scaling (if appropriate). ***"Growth" refers to a slow and steady increase of the rollout of your intervention. Whereas "Scale" is referring to the point where your impact starts to accelerate dramatically.***

9.1. RCT or Quasi-Experimental Design?

Lorem

9.2. Growth

We've [tested](#) and [evaluated](#) our intervention, and now we're ready to start rolling it out. At this point, your goal is likely to grow your organisation to start getting your intervention out there. This roadmap doesn't talk too much about the "behind-the-scenes" of nonprofits, just about the path to impact. So I'm going to recommend the book "[Managing to Change the World](#)" at this point, along with the resources and templates they provide on their website. Especially if a big part of your next step is building organisational capacity for impact growth.

For us in this roadmap, we're going to focus back on Impact, so we need to go back to our [Theory of Change](#) and our [Impact Monitoring](#). We essentially want to update our Theory of Change to reflect the new understanding of our intervention (this may be as simple as taking the [Program Theory](#) you developed for your [pilot](#) and using that, but it's likely you'll need to make a tweak here and there. It often makes sense to have a detailed "**internal**" overall Theory of Change for your intervention, as well as a more streamlined "**external**" Theory of Change, for sharing with interested parties, such as donors). From here, we want to make sure that our [Monitoring & Evaluation](#) processes are in place and robust, and accurately reflect the [Theory of Change](#). From here, we let the ship take its course, learning and improving as we go.

This was another interesting section for me to write, as it feels like one of the more important ones, but it's surprisingly short, because the concepts we need to fall back on have already been detailed (and I don't want to repeat myself). We need to revisit our [Theory of Change](#), and [M&E systems](#), to the point where we've developed a "**learning organisation**". A learning

organisation is a term by [Social Impact Navigator](#), asks you to work through the following checklist:

- ☐ *In our organisation we take time to reflect on our work and its results (using the findings of the social impact analysis as a basis), and discuss this subject. Openly.*
- ☐ *We use the findings of the social impact analysis to learn lessons. We accept that we'll make mistakes, but we use this as an opportunity to learn lessons.*
- ☐ *The organisation's management supports learning and learning processes, and provides incentives for learning.*
- ☐ *Learning processes are an integral part of our work procedures.*
- ☐ *Responsibilities for learning and knowledge management have been clearly allocated.*
- ☐ *There are specific processes for recording results and sharing knowledge.*
- ☐ *We have sufficient resources for the learning process.*
- ☐ *We use available opportunities to exchange experiences with and learn from other organisations.*

And for some projects, this might be enough, a slow and steady growth until we reach a sort of "impact equilibrium", likely depending on your cause area being constrained in some way (i.e. talent or funding constrained), or even just because you work on an issue that is neglected and tractable, but not particularly scaleable. However, for those organisations working on a potentially scalable solution (for example, by scaling up operations in one country, or attempting to explore replication in other countries), it's time to think about our scale strategy.

10. Scale

The Scaling sections here follow the path laid out in Spring Impact's open-source "[Scaling Impact Toolkit](#)".

10.1. Scale Strategy

Now we venture into uncharted territory for me. I'm not at the point of scaling my organisation, but this is supposed to be a Roadmap, and I want to pave a path for when we're ready. Luckily, [Spring Impact's open-source Toolkit](#) is here to help (which I really love, it's a real practical implementation of a lot of theory, which is what I've tried to do with the rest of the Roadmap).

This stage of the Roadmap is unusual, as these following Scaling sections don't intersect with any other [Programme Development Methodologies](#), they are all taken from this toolkit (whereas the rest of the Roadmap is a sort of amalgamation of multiple methodologies into

a cohesive whole). So if you'd prefer, it might make sense to just use their online toolkit (or otherwise use [their services](#), for example, by signing up to their accelerator program!) rather than continue following through this roadmap, as the steps are pretty much the same.

The only place we diverge is that the discussion of the following stages within this roadmap assumes you've followed the steps this far, so it refers back to previous stages to help complete steps here (whereas the Spring Impact Toolkit doesn't do this, so often asks you to create its documents from scratch).

10.1.1. Scale Readiness

This section is based on Spring Impacts' ["Scale Readiness Diagnostic"](#)

Your starting point for scale is a reflection exercise that asks: "Are we ready to scale?"

This can be further divided into two questions

1. How ready is our **Solution*** to be taken to scale?
2. How ready is our **Organisation** to lead that process?

**Solution encompassing your Product, Programme, Service or set of Principles.*

The readiness toolkit asks the following questions in order to assess this (I've only included the questions below, further context and detailed multiple-choice answers can be found at the link above). For each I've tried to note where we should have answers to our questions throughout the roadmap (and if we don't, where we can refer to, to work on this aspect of our solution/organisation).

Solution Readiness

- Is there a clear societal problem that you are committed to addressing and has this been validated by those experiencing the problem?
- Are you clear on who your target end users are and is your solution informed by their needs?
- Do you have a clear understanding of how your solution contributes to your desired social impact (i.e. a Theory of Change) and has this impact been proven?
- Do your end-users value your solution and engage with it over time?
- Are you clear on who will be the payer of your solution once scaled and do you have evidence to support this assumption?
- Is there an ambitious yet realistic scale strategy (i.e., plan for how your organisation will achieve scale in the long term)?

Organisational Readiness

- Is there a clear scale lead with the relevant skills, capacity and dedication to deliver on the scaling plans?
- Does everyone on your team – from staff to the board – support and buy into your scaling ambitions?
- Does your organisation have a culture of testing, learning and using data to validate and inform decision making?
- Is there dedicated leadership support to remove barriers, lead change management, and drive forward your scaling efforts?
- Are the operational functions, processes and systems necessary for delivering your scale strategy and solution well-defined?
- Do you have the financial resources to deliver and iterate on your scaling plans and is this underpinned by a sustainable business model?
- Is your organisation stable and resilient enough to commit to a long-term, challenging scale-up journey?

Addressing Readiness

Based on your answers to the questions above, you should have a clearer sense of whether you're ready to scale and if not, develop clear **Action Points** in order to get there. You may want to consider:

- What are the contingencies? Does the success of anything depend on something having already been completed?
- What elements will come together as you develop your scaling plans, and what needs to be in place beforehand?
- Achieving buy-in from your team is a vital prerequisite before you can successfully scale

From here, determine how you will address each area, (Spring Impact recommends, as do we) using a [RACI framework](#), i.e. for each Action Point, assigning someone to be:

Responsible (for delivering on the action point), **Accountable** (for successfully addressing it), **Consulted** (for advice or input) and **Informed** (of progress). We used RACI within our [Implementation Plans](#) and I like the RACI framework a lot, but you may already have strong processes in place at your organisation for delivering results (like OKRs for example), in which case, I assume that it's entirely possible to address your priority points within your existing system.

Either way, from here, go forth and deliver!

10.1.2. Problem Definition

This section is based on Spring Impacts "[Problem Definition](#)"

Hopefully you're getting [déjà vu](#) at this point - didn't we do this as part of our intervention development? That's right! Great spot 😊

I still think it makes sense to review (and potentially update) your problem definition at this point though, in case anything has changed since you completed your [Pilot](#). Hopefully, it should be relatively easy though. Spring Impact asks you to do this by identifying a) the **Root Causes**, b) the **Problem Itself**, c) **Those Affected**, and d) **The Implications**. Again, hopefully your previous [Problem Definition](#) (and its sub-sections) already contains the answers to these questions.

Some questions you can ask yourself to refine your problem definition are:

- Is there a big gap between what is available today and what is needed to address the problem?
- Are you sure you've identified the main problem rather than a symptom of the problem?
- Does it avoid referencing any solution? Make sure you're focusing on defining the problem, not how it's solved.
- If you identify lots of problems, is there anything that unites them? Could they all be roots of a single problem, or symptoms of a single problem?
- Is the problem you are solving unique? How is the problem you're addressing different from what others are focusing on?

10.1.3. Intended Impact

This section is based on Spring Impacts "[Intended Impact](#)"

Let's move on to outlining your intended impact, starting with answering the following prompt:

"How will the world be different in 10 years if I succeed in addressing the societal problem?"

Additionally this section is pretty similar to our [Project Objectives](#) we did when creating our [Theory of Change](#). As with the previous section, it probably does make sense to revisit this at this point as a sense-check, but probably no need to reinvent the wheel.

From here, we work on the specifics and figure out where the impact is happening both at the intended change level (i.e. individual, organisational or societal) and geographically. In order to do this, we can work through the following questions:

- Who is the target constituent? What is the geographic scope?
- How big is the target constituency within this geography? (estimate)
- What is success? What does impact look like?
- What % of the target constituency do we want to impact?

You likely already have answers to quite a few of these, probably from your [Evidence Review](#) or [Needs Assessment](#).

10.1.4. Design your Core

This section is based on Spring Impacts "[Design Your Core](#)"

At this point, we're trying to decide what exactly it is that we're scaling. Is it a Product? A Programme? A Service? A set of Principles?

From here, you design your core. This means understanding which of your 'ingredients' are **Essential** (elements that are absolutely necessary to achieve impact), and which are **Non-Essential** (elements that are nice to have). Examples of ingredients are:

- Values that underpin your solution;
- Processes that support your solution;
- Key messages or tactics (i.e the social action activities used to achieve a campaign's goals).

After this initial sweep, you should challenge yourself to consider why these ingredients are essential (the more complex it becomes, the less scalable it likely will be). Of the essential ingredients, can any be flexible/adaptable (i.e. if you are scaling a campaign, the message of the campaign must be exact to create its intended impact, but how local campaigners choose to deliver the messages can be flexible – be it through a public protest or through handing out pamphlets.). You should now be able to create the recipe for your core, with ingredients divided into **Essential**, **Flexible** and **Non-Essential**.

10.1.5. Identify Risky Assumptions

This section is based on Spring Impacts "[Identify Risky Assumptions](#)"

Hopefully more [déjà vu](#) for us, which is good! Scaling isn't as scary as we thought 😊. Again here, the toolkit provided by Spring Impact is slightly different to our previous method of doing this (using [Programme Theories](#) to tease them out, among other methods), and I think it probably makes sense to use what we've already done as a starting point, but also to follow this new template, as a way to revisit and sense-check our previous assumptions.

To start with then, start [brainstorming](#) assumptions. For example:

- What parts of your solution, strategy, or scale model are untested?
- What are your biggest unknowns?
- What feels like the biggest risks to your success?

As with [Identifying Assumptions](#) in our Theory of Change section, break down every assumption to be as concise as it can be, such as breaking down 'Users will pay for this model', you can break this down into 'Users are interested in the model' and 'Users will have the means to pay for it'.

Now it's time to prioritise testing. We can use a matrix to help here, with the following axis:

- **High-Risk to Low-Risk** (with High-Risk having the biggest impact if they fail, and Low-Risk having the least impact)
- **Certain to Uncertain** (with Certain being the ones you are most confident about due to previous data or learnings, and Uncertain being the ones you have least confidence in)

Then depending on where your assumptions fall in the matrix, they can be categorised as follows:

- **High-Risk / Uncertain** - Test first so we don't go down the wrong pathway
- **Low-Risk / Uncertain** - Keep an eye out for
- **High-Risk / Certain** - Accept the associate risk and cope
- **Low-Risk / Certain** - Safely deprioritise

As you might have guessed, it's the High-Risk / Uncertain quadrant that we're going to test next...

10.1.6. Test Risky Assumptions

This section is based on Spring Impacts "[Test Risky Assumptions](#)"

You can either reuse [Implementation Plan](#) templates you developed during the [Viability](#) / [Impact](#) Testing phases of the roadmap, or use the template that Spring Impact provides. Again, I don't think there's a right answer here (I think I'm inclined to use the [Implementation Plan](#) because it allows you to go into more detail, and have everything in the same place). But rather than reiterate the Implementation Plan advice from earlier, I'll outline the Spring Impact template instead (it's important to note that you create one of these for each assumption):

- **The Assumption** - What assumption are you testing?
- **Objectives** - What do you hope to achieve and/or learn?
- **Type of test** - What type of test will you use?
- **Who** - Who do you need to engage?
- **What** - What are you asking participants to do to participate in the test? What do you want participants to do as a result of the test?
- **Success metrics** - What does success look like?

- **Data Needed** - What quantitative and qualitative data will help you test your assumption?
- **Action Plan** - What are the steps of your test and how much time is needed?

10.1.7. Sweet Spot for Scale

This section is based on Spring Impacts [“Sweet Spot for Scale”](#)

The “sweet spot for scale” refers to the intersection of a Venn Diagram for **Impact**, **Value** and **Sustainability**. The purpose of this section, is to analyse the solution for each of the three lenses:

- **Impact - *does it create a lasting change?***
 - Does the solution meet its intended impact and actually make people's lives better?
 - Does it make a meaningful and equitable contribution to addressing the societal problem at scale?
 - Has this been validated?
- **Value - *will people want it and stick with it?***
 - Do people actually want and demand the solution?
 - Do they engage with it over time?
- **Sustainability - *can it impact a lot of people sustainably?***
 - You need to understand what the costs are and where the resources are going to come from to cover these costs.
- **Overall**
 - Where do the biggest challenges arise for you in finding the sweet spot for scale?
 - What do you need to prioritise improving in order to improve your scale readiness?
 - What are some of the barriers you might face when working to improve your scalability? How might you overcome these challenges?

10.1.8. Measure what Matters

This section is based on Spring Impacts [“Measure what Matters”](#)

Back on familiar ground, this is close to our [Impact Monitoring](#) systems we have in place. But Spring Impact's [Critical Metrics Calculator](#) is a pretty neat tool (it's effectively a [Cost-Effectiveness Analysis](#)), and as we know from our discussion about [meta-tools](#), it often helps to use multiple tools to see if they converge on an answer.

The calculator is divided into two overall sections: “Impact and Value”, and “Rough Cost”, using the Goal you defined in your [Intended Impact](#) as a starting point.



Impact and Value

In this section, you put in numbers such as:

- **Target Constituents** (the number of people impacted by the societal problem)
- **% of total reached** (the % of all target constituents that know your solution exists)
- **Conversion rate** (% who know about your solution and choose to participate or sign up)
- **Engagement rate** (% of people who have signed up, who stick with and fully engage with your solution)
- **Impact rate** (% who achieve a successful outcome)

Which then gives you numbers such as:

- **# constituents signed up**
- **# constituents engaged**
- **# constituents impacted**

Rough Cost

In this section, you fill in the **per unit cost** and **quantity** for costs such as:

- **Biggest cost driver 1** (Example: High level staff)
- **Biggest cost driver 2** (Example: Mid level staff)
- **Biggest cost driver 2** (Example: Junior staff)
- **Cost Per Acquisition** (cost to reach new target constituents)
- **Fixed cost estimate** (include fixed costs such as set up costs if applicable, delete this row if not)
- **Admin/Misc**

Which then gives you numbers such as:

- **Total cost**
- **Cost per person signed up**
- **Cost per person engaged**
- **Cost per person impacted**

10.1.9. Doer and Payer

This section is based on Spring Impacts "[Doer and Payer at Scale](#)"

Here, my write-up differentiates slightly from Spring Impact's "Doer and Payer at Scale" section, as that is itself inspired by The Stanford Social Innovation Review (SSIR), which I've

used as the basis instead. The SSIR [asks you to answer](#) the following questions when determining your model to scale:

1. Will your model scale?

- a. **Is it cheap enough?** - You should have a good idea of this based on your [Cost-Effectiveness Analysis](#) (which can be updated after your Pilot).
- b. **Is it simple enough?** - Again, you should have a good sense of this based on an updated [Theory of Change](#) after your Pilot. The more complicated the model the less likely you'll be able to have others replicate it.
- c. **Is it adaptable enough?** - Finally, the data you captured as part of your [Process Evaluation](#) can help with this (as well as data captured using the 'Transportable' CART Principle). You should have a good sense of what works about your intervention and why (and therefore, where the uncertainties lay around new contexts).

2. Who's going to do it?

- a. **Your organisation** - This gives you the most control over the intervention, however building your organisation to scale can be difficult and has drawbacks.
- b. **Other NGOs** - If you're able to have your model replicated at other NGOs (i.e. think of examples like the [Open Wing Alliance](#)) then this can really ramp up your scaling to impact. However, this can be hard to pull off [for various reasons](#) (other NGOs wanting to be unique, or implementing a poorer version of the intervention etc.).
- c. **Other Businesses** - In particular, this makes sense if your solution can be a profitable business model/market for an industry (just one or two existing businesses might not make sense to scale). This can work for more social enterprises.
- d. **Governments** - Huge bandwidth and resources. However can be inefficient and inconsistent, so only really works for delivering basic service solutions at a huge scale.

3. Whos' going to pay for it?

- a. **Customers** - i.e. likely through a social enterprise model. Some non-profits do make some money selling merch etc. but that's not really what we're looking at here.
- b. **Taxes** - This is the obvious route if your doer is a government. Additionally, this model doesn't necessarily require raising taxes, it could just be a redistribution of current tax.
- c. **Big Aid** - This is referring to revenue typically from governments, rather than wealthy philanthropists. In some cases the government will want to be the doer as well, in other cases, they pay another doer.

- d. **Private Philanthropy** - And finally, this includes any and all donations from private philanthropists, including big donors like the Bill and Melinda Gates Foundation and Open Philanthropy, down to individual small donors.

At different scales your “doer’s” and your “payer’s” may be different. The best place to start is understanding who your doer and payer is for your current number of beneficiaries (the answers to these questions at this point are likely “You” and “Private Philanthropy”).

From here, imagine your reaching a number of beneficiaries 10x that of your current reach. What does the Doer/Payer relationship look like here?. What about 100x? 1,000x? Your Doers and Payer will likely change at these different orders of magnitude (likely with some changes to your model along the way), and thinking through this exercise will give you a sense of the outline of your strategy to scale.

It’s a useful concrete step to summarise your choice of doer and payer at different levels. For example (your examples may not be as strikingly dissimilar as the examples below, you may have the same Doers and Payers at multiple levels):

- **At 10x** - to ensure widespread social sector adoption, we need to make sure that our solution is as simple and replicable as possible.
- **At 100x** - to get the commercial sector, we should show that this can produce a profit margin
- **At 1,000x** - we want government to pay for this eventually, so cost effectiveness is a big priority

10.2. Scale Model

Lorem ipsum dolor sit amet

10.3. Scale Implementation

Lorem ipsum dolor sit amet

Appendix

Programme Development Methodologies

Here are a number of methodologies we drew on to create our Impact Roadmap (with the names and links of the methodology on the left, and where they influenced this Roadmap on the right):

- [IDInsight: Impact Measurement Guide](#) - Overall Structure
- [PHINEO: Social Impact Navigator](#) - Overall Structure
- [NPC: Theory of Change in Ten Steps](#) - Theory of Change
- [SparkWave: Ten Conditions for Change](#) - Intervention Development: Change Mechanisms
- [PubMed: 6SQuID](#) - Intervention Development, Viability Testing, and Impact Testing
- [INTRAC: Outcome Mapping](#) - Viability Testing: Potential Deliverables
- [IPA: Right-Fit Evidence \(CART Principles\)](#) - Impact Monitoring: The CART Principles
- [OECD: EvalNet](#) - Impact Testing: Process Evaluation
- [INTRAC: the M&E Universe](#) - Impact Analysis: Monitoring & Evaluation
- [SpringImpact: The Scaling Impact Toolkit](#) - From Growth to Scale
- [The Goldilocks Challenge](#) - Impact Monitoring
- [Managing to Change the World](#) - Not used in the Impact Roadmap, but a very useful book filled with practical advice and templates for managing your nonprofit organisation

Brainstorming

The following ideas are taken from [The Thinking Shop](#) and are available under a Creative Commons licence

Limitation

Write a Clear Brief - Failing to provide a clear and compelling brief does not, contrary to popular opinion, 'open up' creativity. Use [this briefing template](#) or at least provide a clear objective in the form of a single simple sentence.

Generation

- **The Meta Sequence**
 - **Deep** – *Research & Define*
 - The more information you expose yourself to, the more your brain has to work with. Deep dive before an ideation session and define the objective.
 - **Wide** – *Creative & Divergent*
 - Try to minimise critical evaluation while brainstorming and exploring new ideas. Creative insights are helped by a mindset of playful curiosity.
 - **Narrow** – *Critical & Convergent*
 - Use objective evaluation criteria where possible. The best ideas are vulnerable as they don't conform to expectations; so don't use critical

analysis to leave only the safest ideas standing, but to find out which challenging ideas might actually be viable.

- **Up** – *Synthesise & Iterate*
 - Once you have an idea that works, it's tempting to cling to it. However, the best ideas often come from having the curiosity and courage to push beyond what works to something more interesting. Keep playing to find a deeper synthesis.

10 Ideas in 10 Minutes - Force your internal critic to shut up by generating as many ideas as possible. Quantity over quality will help your mind to expand rather than contract. Set a stopwatch for 10 minutes to write down 10 ideas. Importantly, they don't have to be good ideas.

Draw on Inspiration - Instead of writing ideas, simply start sketching relevant concepts. By activating your visual cortex you'll open up new ways to think about things. Don't try to solve the problem right away, just start moving the pencil.

Shower Thoughts - The reason we have great ideas in the shower is that we're not doing anything else. Use your time in the shower, on the train, during your run, or meditating in a sensory deprivation chamber to consider the brief.

101 Ideas - Give yourself or your team a deadline to compile 101 ideas in a list. By having so many ideas you give yourself permission to have more bad ideas, which in turn opens you up to more good ideas.

Go Slow - Immerse yourself in research and visual references, then let your subconscious go to work. Sleep on it and give your mind time to synthesise. Try this before you start and/or between sessions.

Tricktionary - Use random stimuli like reading out words from the dictionary to provoke a response, or try using the last letter of one word as the starting letter for another word. Free-associate words and concepts to spark connections.

Mind Map - Write your objective in a bubble in the middle of the page and branch out nested and associated ideas to discover new perspectives as stimuli. Try putting your mind map on the wall while brainstorming.

Perspective Shifts

Inversion - By flipping the perspective you can open up new insights. What's the opposite of the problem or solution? What would be the worst idea? Who isn't the target audience?

Recontext - How would you solve the problem in the future or past? What about a different cultural context? How would aliens see things? Try 5 ideas in 5 minutes to generate different perspectives then pick 2 or 3 to ideate from.

No Limits - Removing limitations means removing assumptions, which can result in novel perspectives and solutions. Imagine you have unlimited superpowers, budget, time, or technological capability to solve the problem. What would a god do?

Extremify - Looking at extremes can bring insights into sharp relief. What would happen if you amplified the problem or solution? What if it was minuscule or all-consuming; what would a global response look like?

Challenge - Creative breakthroughs often arise from challenging conventions. What assumptions have already been made? Could they be wrong, or shifted? What if the brief is too narrow? Too wide? Pretend to be your competitor.

- **Imagine You're...**

- **The Devil** - What would someone ruthless do? How could you use mischief to disrupt the situation? What if we broke some stuff? How might someone evil exploit you?
- **The Customer** - What does a day in the life of your target audience look like? What are their fears and aspirations? What might they not understand? What might we not understand about them?
- **The Sage** - Take on the persona of an oracle. Why are we really doing this? What's the true insight and purpose of this project? Is there a deeper truth? What would someone truly wise do?
- **The Explorer** - How can we break new ground? Are we thinking too locally? If we could rewrite the rules, what would we do differently? How might someone from a different culture approach this?
- **The Innocent** - What would you think if you knew nothing about this? How might a child interpret the problem or solution? What's the obvious solution and why is it good or bad?
- **The Joker** - Imagine you're a famous comedian who fits the tone of the brief. What would they joke about? What's an irreverent insight? Is there an elephant in the room that needs to be outed?
- **The Ruler** - What would you do if you were the boss, a wealthy investor, or someone with absolute power? If you already had the support of everyone, how might you do things differently? What might a judicious leader suggest?
- **An Alien** - What would an alien make of this brief? By writing from the perspective of an extraterrestrial we can discover insights and assumptions to

help us think in a more 'first principles' way. Check out the Strange Planet comics online for inspiration for alien perspectives.

- **The Disruptor** - Pretend you're an entrepreneur who is aiming to shake up the industry or status quo with a new vision for what might be. What would a disruptor do to challenge the existing order with a new take on things? How could you forge a new path?
- **The Artist** - How might different kinds of artists interpret this problem? Brainstorm from the perspective of a visual artist, thespian, fashion designer, advertising creative, product designer, UX designer, or poet.

Another's shoes - Meta-brainstorm different perspectives and archetypes that might be relevant for your particular situation. For example, if you have a marketing problem you might suggest different kinds of customers and business stakeholders. Then generate ideas from their perspectives.

Mental Models

The Pareto Principle - About 80% of the output tends to come from 20% of the input. How might we optimise by focussing on the most relevant factors? How can we design for the primary audience or most valuable outcome?

First Principles - Define the base principles to reason more clearly. Question what's actually fundamental to bring the problem (and solution) into sharper relief. What unnecessary assumptions are being made? What if we had to make everything from scratch?

Social Proof - As tribal creatures, we want validation before we act. How can we create and communicate social buy-in? Could the crowd itself contribute in some way? Is there a way to provide a social feedback loop, testimonials, or other social validation?

Critical Mass - A critical mass is a self-sustaining chain reaction. What could we do to create a viral effect whereby the idea spreads itself? How might we make it remarkable or useful enough to become its own amplifier?

Scarcity - We value things that are in short supply. Is there a way to create demand or at least the perception of it? How might limitations be beneficial? If there isn't a ready-made market or audience can we create one?

The Third Story - Consider the impartial perspective. We're often blind to our own assumptions, and our perceptions are coloured by our insider understanding. What might someone with no knowledge or bias think?

Second-Order Thinking - Consider the flow-on effects: what might the consequences of the problem or solution entail? What are the variables at play? How might the family of your target audience react? How could time, complications, or feedback loops have influence?

Provocation

Future News - Pretend you've already achieved success. What would the newspaper headline or article say? This will spur you to think of big, newsworthy ideas. Explaining how and why it worked so well forces you to consider the steps to success.

Minimum Viable Idea - The creative process is iterative, and by getting to the minimum it can open us up to pivot in relevant and interesting ways. What's the minimum needed to achieve the objective? What's the most efficient, quickest, and cheapest option you can think of?

Five Whys - This technique is often used to get to the root of a problem, but it can also discover insights that lead to innovative thinking. Start with a problem or solution, then keep asking 'why?' to see where it leads you.

Journey Map - Write your problem or starting point at one end and the ideal outcome at the other. Now fill in the gaps to map out what needs to happen to bring your objective into reality. What steps will need to occur? How might they influence each other?

What If? - Simply start the thinking process with 'What if?' can elicit a creative response. Try saying it out loud or writing it down several times in a list to hack your brain to think differently. Build upon other ideas by challenging them with 'what if' scenarios.

SWOT Now - Create four quadrants and list Strengths, Weaknesses, Opportunities and Threats. S&W relate to internal factors while O&T tend to relate to external aspects. SWOT analyses are usually used for strategic evaluation, but they can also be used to provoke new ideas.

Change of Scene - Go outside, to a bar for a few drinks, or to the top of a mountain and see how the change of environment changes your perspective. Many people find walking helpful for ideation and stimulus.

Will it Blend? - How could you combine two or more ideas, techniques, aspects, solutions, or problems? Try writing relevant keywords or ideas down on bits of paper then mashing them up randomly.

See Inspiration - Use the room you're in as a visual inspiration canvas by finding references to print and put up on the walls, along with your rough concepts, sketches, headlines, and so on. Try some image searches for related ideas and save the ones you like as stimuli.

Reverse Brainstorming - Instead of solutions, brainstorm to generate problems. What other issues might need solving? Will there be flow-on effects that require different solutions? What's a different way of thinking about the problem space?

Starburst Questions - Create a mind map star with six points for each of the classic analysis questions: Who, what, when, where, why, and how. Note that the 'why' is the most likely to provide relevant insight, and that the 'how' can sometimes be a distraction.

In Groups

YourBias.Is - Group brainstorming is often affected by social factors and confident voices thwarting more considered thinking. Familiarise yourself with common cognitive biases that can affect decision making, evaluation, and group dynamics at yourbias.is

Brain Writing - Everyone writes an idea on a piece of paper. Pass your piece of paper clockwise, the next person builds on the idea. Repeat until everyone has contributed to each other's ideas.

Vote No to Bias - Our judgments are invariably shaped by our personal biases. So, it's best to have discerning independent third parties evaluate anonymised ideas rather than a group vote.

Diverse Inputs - Creative insights often arise from novel perspectives, so try to facilitate diverse group dynamics e.g. put engineers with designers, management with front-line staff, and so on.

Step Ladder - Everyone brainstorms individually. The first two people present ideas to each other. A third person is invited into the room to present their ideas. Repeat until everyone has presented unaffected by others' ideas. Finally, recap all ideas for the group.

Heads in the Cloud - Rather than verbal collaboration, use a cloud-based service like Google Docs with open sharing permissions so everyone can anonymously collaborate, comment, and build on each other's ideas.

Anonypost - Each person writes or draws a single idea per sticky note and puts it up on the wall in a separate room. Take turns nominating your favourite idea that wasn't your own (and hasn't yet been nominated) and explain why.



Role Up - Give each person in the group a role to play such as the customer, the rebel, the narrator, the joker, the sage, the CEO, the shareholder, the lunatic, etc. Then give a scenario involving the product, problem, or situation to role play

Concept Wall - To start your collaborative brainstorming session, everyone writes or draws a few key concepts relating to the subject matter on sticky notes. Then put them up on the wall as a stimulus for further brainstorming.