BIRDLAB
BEHAVIORAL INSIGHTS RESEARCH AND DESIGN

unicef
for every child

DEPTHS

An Applied Behavioural Science Field Guide

Practical Tools for Stronger Programmes

TABLE OF CONTENTS

1. **Introduction**
2. **Define**
3. **Explore & Diagnose**
4. **Prototype Designs**
5. **Test Hypotheses**
6. **Scale**
7. **Conclusion**

Introduction

Welcome!

Applied behavioural science has become an essential capability for UNICEF and the wider development and humanitarian sectors. It strengthens programmes and policies by helping them address the behavioural and social factors that shape families' and children's well-being.

This guide lays out an evidence base and methods for UNICEF's approach to applying behavioural science, and shows how the approach works in practice through concrete examples.

At a glance

1. Introducing the Field Guide

This section explains the purpose of the Field Guide, its intended audience, and how the chapters are structured, helping readers to quickly locate what they need.

2. Applied behavioural science

This section introduces the field of applied behavioural science and what lessons it provides about how people think and act in context.

3. UNICEF's approach: DEPTHS

The section gives an overview of DEPTHS: a process to help practitioners apply, manage, and advocate for behavioural science in their work.

Introducing the Field Guide

Behavioural science is the study of how people make decisions and take action in their daily lives. At UNICEF, it is increasingly used to design evidence-informed approaches that support children's and families' opportunities to enhance their well-being.

Crucially, behavioural science helps to uncover the barriers that prevent people — even when informed and motivated — from turning intention into action. Rather than a single discipline, behavioural science is a multidisciplinary practice, bringing together evidence and insights from fields such as social psychology, economics, and cognitive neuroscience to better understand and influence behaviour.

Development and humanitarian organizations have embraced behavioural science approaches at remarkable speed over the last decade. Nearly all major international organizations now use behavioural insights in some part of their work. Within the UN system, behavioural science is recognized as one of the five core capabilities needed to deliver on the Secretary-General's UN 2.0 vision, and the UN Behavioural Science Group now links more than 1,000 members across over 60 entities, including UNICEF, WHO, and UNDP. The World Bank's eMBeD unit has documented this global uptake, profiling the rapid expansion of behavioural approaches across agencies like Save the Children and the International Rescue Committee.

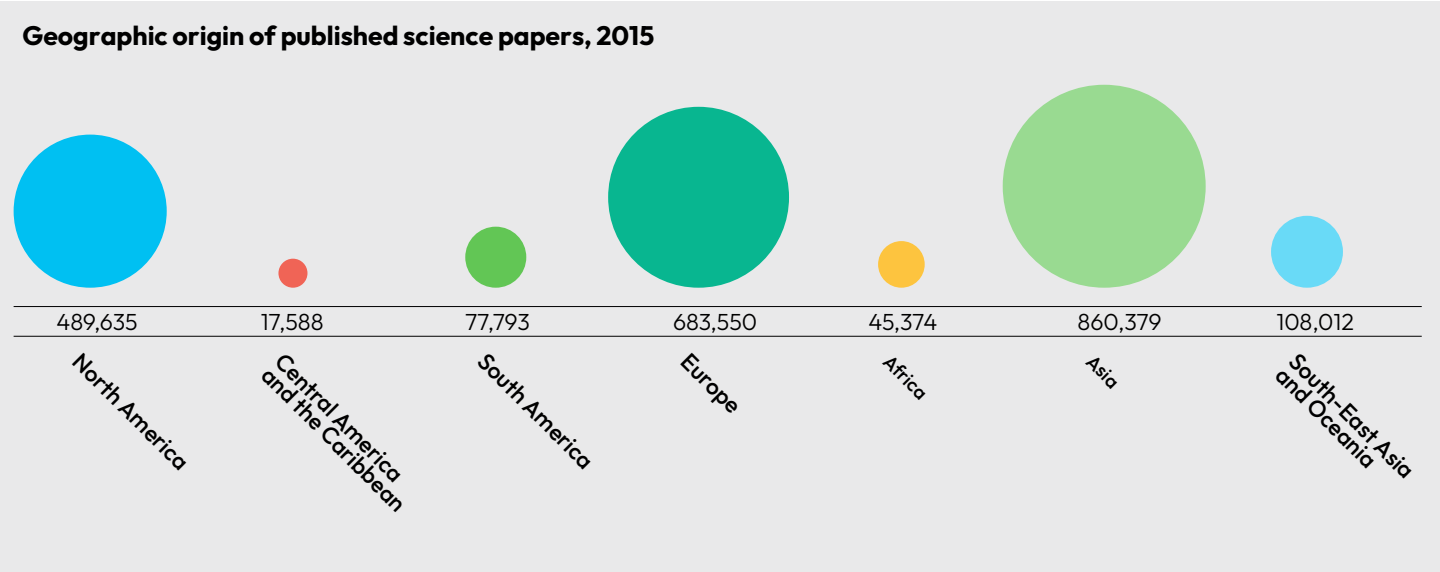
These approaches are already shifting practice on the ground. Behavioural insights have informed efforts to increase tax compliance, improve reporting of violence in Indonesia, and expand access to family planning. They underpin work to promote handwashing among schoolchildren in the Philippines and strengthen childhood malnutrition screening in Mali. In humanitarian settings, behavioural science is beginning to demonstrate its value as well—helping to engage refugee parents in early childhood development and improving medical incident reporting in camps in Bangladesh and Rwanda.

Yet adoption across the sector remains uneven. Some organizations are experimenting cautiously, while others—including UNICEF's Behavioural Insights Research and Design (BIRD) Lab—have established teams and growing portfolios. Even within UNICEF, adoption varies widely: some teams apply behavioural science methods routinely, while others are just starting out. These inconsistencies mirror a broader challenge in global evidence generation. Most behavioural science research relies on participants from Western, educated, industrialized, rich, democratic (WEIRD) societies, who account for a disproportionate share of study samples despite representing a small fraction of the world's population[1]. This gap limits the relevance of existing evidence for many of the communities UNICEF serves.

The American Psychological Association found that people from Western, educated, industrialized, rich, and democratic (WEIRD) societies make up as much as 80% of study participants, yet represent only 12% of the world's population. They are not only unrepresentative of humanity as a whole, but on many measures, they are statistical outliers.[2]

As UNICEF works with diverse populations, this field guide aims to promote more consistent and effective use of behavioural science, highlight where it can add value, and contribute to building a more inclusive and representative evidence base. The examples, case studies, and worksheets that follow are meant to inspire and guide — but they should always be adapted with care, informed by local perspectives, and tested in context.

1 https://www.apa.org/monitor/2010/05/weird
2 https://www.apa.org/monitor/2010/05/weird



Data via worldmapper.org CC BY-NC-SA 4.0

Who is this guide for?

This resource is for UNICEF, partners, and governments interested in applying behavioural science evidence and methods to advance the SDGs and secure results and rights for children. It is designed to meet the needs of three primary audiences:

- **Managers** of teams and projects who want to leverage applied behavioural science to improve programme or policy outcomes
- **Practitioners** who are directly applying behavioural science, often with prior training in related areas of Social and Behaviour Change (SBC)
- **Leaders** who are advocating for more systematic and evidence-based approaches to behaviour change

What to expect

The introductory chapter provides a summary of applied behavioural science. Each subsequent chapter in the DEPTHS process begins with an overview, outlining the chapter's focus (e.g. tools and processes), what teams will do during each phase, and common challenges.

Each chapter offers practical guidance on when and how to use specific behavioural science tools. These tools can be applied on their own or integrated into existing SBC processes to strengthen them. Each section also provides curated links to additional resources for further learning.

Related resources

The field guide is part of a comprehensive set of resources that aims to serve various audiences and use cases. Readers and teams using the field guide will explore key concepts and see how applied behavioural science works in practice, while experienced practitioners can dive deeper into tools and methods to strengthen their skills. The full package includes:

- **Field Guide:** The core reference or "textbook." In this package, the field guide brings together tools, case studies, templates, and techniques to support the application of behavioural science from start to finish.
- **Toolkit:** Worksheets and tools for each phase of DEPTHS, illustrated with use cases and examples to support understanding.
- **Training Materials:** Participatory lecture modules with practical exercises that help practitioners and partners build their skills.
- **Microsite:** Housed on the BIRD Lab website, the microsite offers the most up-to-date materials and is updated regularly as the package evolves.

Materials will be available in English, French and Spanish, with additional translations developed based on demand.

Applied behavioural science

Behavioural science is the empirical study of how people think, decide, and act, with an emphasis on observation and experimentation.[3] While it originated at the intersection of economics and cognitive psychology (known as behavioural economics), it has since broadened to include research methods, evidence, and insights from many other disciplines.

This field guide draws on three approaches: behavioural science, human-centred design and systems thinking, which together inform the DEPTHS process by explaining how people make decisions. It highlights the systems that shape behaviour, and grounds solutions in lived experience. What distinguishes behavioural science as a subdiscipline of the social sciences more broadly is its focus on two elements: first, its particular focus on the interaction between cognition and context — i.e. how local, often changeable details in the environment affect people's decisions. Second, its commitment to rigorous evidence, building knowledge about human behaviour through systematic experimentation and evaluation.

Behavioural science has its roots in academic research, but its full potential is realized when it is applied in practice. Applied behavioural science builds the evidence and knowledge of behavioural science to intentionally influence people's behaviours, habits, and decisions in real-world contexts. At UNICEF, this work is especially focused on improving opportunities for children, along with their health, well-being, and safety.

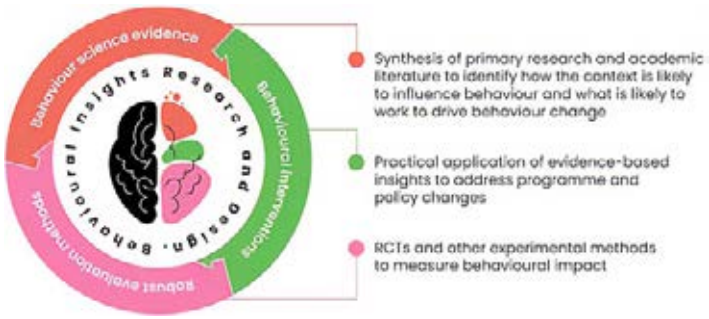


Figure 2: Behavioural Insights Research and Design.
Source: Busara, 2025 based on BIT.

As described in Figure 2, applied behavioural science offers practitioners the following:

- An evidence base, which includes a conceptual model of the mind, a lens to analyze situations, and practical guidance on approaches that may be effective in a context.
- A process informed by that evidence, which enables the diagnosis of the drivers of specific decisions and behaviours, which can be influenced with targeted interventions.
- Robust evaluation methods, which assess the impact of interventions and generate new lessons to inform the evidence base.

Behavioural science recognizes that human decision-making is often shaped by mental shortcuts that help people to make quick decisions, but can also lead to predictable errors. For example, availability bias leads to overestimating the likelihood of events that are easy to recall, such as plane crashes or shark attacks, because they feel more salient than less memorable risks. While these biases and heuristics are important, they are well covered by resources such as UNICEF's The Behavioural Drivers Model and summaries by groups like the Decision Lab. There is also more information on biases in these training resources: Lesson 1, Lesson 2, and Lesson 3.

Rather than revisiting underlying psychological mechanisms, the focus here is on operationalising insights that researchers have uncovered about the drivers of behaviour and decision-making. The field guide is designed to help readers apply behavioural science methods to better understand and influence the contexts and communities they work with.

The Behavioural Lens

A way of viewing challenges through evidence of how people actually think and act, the Behavioural Lens helps to reveal barriers, drivers, and opportunities for change. At its core, this perspective is grounded in understanding the different mental processes that interact with our environment and shape our behaviour.

3 UNICEF BIRD Lab

Dual Process Theory

One of the most useful metaphors for decision-making is Dual Process Theory, popularized by the Nobel Prize-winning psychologist Daniel Kahneman. It explains that our minds rely on two systems: a fast, automatic, and impulsive one ("System 1"), and a slower, more deliberate one ("System 2").

System 1, the reactive system, is essential to our survival. In fact, many of our cognitive functions rely on mental processes that are "fast, effortless, and automatic" — such as recognizing faces or triggering the fight-or-flight response. Imagine if one had to slow down and consciously analyze the facial features of each person we encountered, and consciously compare them to memories of previous faces we've seen? The effort would be overwhelming and leave little capacity for the many other tasks that occupy our conscious mind daily.

System 1 serves us well most of the time, but its speed and efficiency comes with trade-offs. Its automatic nature and susceptibility to environmental cues can make it prone to error, especially in situations that require careful judgment. In short, System 1 relies on mental shortcuts to deliver fast answers, but those answers can often be incomplete, misleading, or biased.

In contrast, System 2 thinking refers to our more deliberate and conscious thinking. System 2 processing is slow, reflective, and attentive to explicit goals and intentions.[4] By pausing, monitoring, and correcting judgments, System 2 can counter the errors that arise from System 1's quick intuitions. However, it doesn't automatically override them — engaging System 2 requires consciously slowing down, paying attention, and reflecting.

Even so, System 2 is imperfect. Logical thought is not like a calculator — people often struggle with complex equations or probabilities, even when thinking carefully. System 2 thinking also comes at a cost: Effortful focus reduces our capacity to plan or think about other things, and sustained use can leave us tired or drained. When overloaded, this kind of thinking becomes less effective.

A useful metaphor is the rider and the elephant. System 2 is the rider — conscious, deliberate, and capable of direction — while System 1 is the elephant — powerful, fast, and doing most of the work.[5] The rider can sometimes steer, but the elephant is often in charge. In fact, our conscious minds frequently invent rationales for actions that were actually driven by System 1.[6]

Figure 3: The rider (System 2) and the Elephant (System 1).

A narrative of the mind and the opportunity for change

Building on Dual Process Theory, four core insights from behavioural science can be organized into a narrative that interprets people's decisions and behaviours:

- **Mental limits:** People have limited time and energy, and most of their mental effort goes into a small number of complex decisions each day.
- **Shortcuts:** For everything else, people rely on mental shortcuts: looking at what others are doing, following what feels good, or defaulting to familiar routines. One of the most powerful shortcuts is habit — repetitive actions performed with little conscious thought, like cycling, eating, or checking a phone.
- **Imperfect outcomes:** These shortcuts work most of the time but not always; they can lead people to act in ways that conflict with their own goals or intentions.
- **Opportunity for change:** By understanding how shortcuts shape behaviour, practitioners can adjust the surrounding context to encourage healthier, safer, or more beneficial choices.

Consider how these four insights play out in a sample situation: a caregiver is deciding whether to vaccinate their child. What steps are they likely to take?

4 Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.
5 See Haidt's The Happiness Hypothesis which popularized the metaphor in modern times. The original metaphor is attributed to the Buddha.
6 For example, see Wilson (2002) with his aptly named book Strangers to Ourselves.

- Do individuals read the full body of scientific literature before deciding on vaccination? Of course not, our cognitive limits make that impossible. Instead, people rely on mental shortcuts: what they already know, stories they've heard, how easy it is to reach the clinic, or whether someone in their community got sick after a shot. People will use these cues — context, social networks, and perceptions of others — to guide their choices.
- However, shortcuts don't always lead to the best outcomes and can often be misleading. Stories of illness may be myths, or unrelated to vaccination; a long wait at the clinic one day may not reflect the usual experience.
- The mind can't avoid using shortcuts, but practitioners can design around them. By reducing barriers and highlighting positive norms, the same shortcuts can be harnessed to help people follow through on their intentions.

There is extensive research on mental shortcuts, how they work, and how to harness or address them. References are provided at the end of the chapter for readers who want to dive deeper. Now, it's time to examine practical examples of these mental shortcuts in action.

Examples of biases or shortcuts

Behavioural scientists have studied over a hundred different shortcuts that the human mind relies on.[7],[8] It isn't necessary to know all of them to apply behavioural science effectively — the research process itself helps identify the specific issues at play in a given situation. Nonetheless, understanding a few of these common shortcuts can be valuable. Below are some examples that are relevant to UNICEF's work.

- **Present bias** is the tendency to give more weight to immediate costs and benefits than to those that occur in the future. It is strongest when benefits are long-term or abstract, while costs are immediate and concrete, and vice versa. [9]

  Consider an SBC programme designed to encourage immunization as an example. The health benefits are delayed — protection against illness at some point in the future — whereas the costs are immediate: travel time, or the inconvenience of a clinic visit. An SBC programme that emphasizes these delayed medical benefits may therefore fall flat. Instead, a more effective approach might highlight a near-term social benefit, such as how many peers or neighbours are already vaccinating their children, making the decision feel more immediate and rewarding.

- **Defaults and status quo bias**. When people are presented with options, the default choice carries tremendous influence. This is partly due to status quo bias: the tendency to stick with existing conditions even when there are strong reasons to choose otherwise.

  Consider a savings programme for employees at a major company in Afghanistan.[10] Researchers tested several strategies to boost enrollment: offering the programme with no incentives, providing a 50% matching contribution, and automatically enrolling employees by default into the programme (with the option to opt out). Enrollment was nearly zero with no incentives or defaults. The 50% match increased participation by 47%, while the default increased it by 40%. Although the match was costly and difficult to sustain, the default was inexpensive and preserved employees' freedom to opt out. When combined, defaults and incentives proved even more effective. This example illustrates the power of a choice architecture: Simple changes to how choices are presented in order to drive a change in behaviour.

7 Blumenthal-Barby J. S. (2016). Biases and Heuristics in Decision Making and Their Impact on Autonomy. The American journal of bioethics, 16(5), 5–15. https://doi.org/10.1080/15265161.2016.1159750

8 Bojke, L., Soares, M., Claxton, K., Colson, A., Fox, A., Jackson, C., Jankovic, D., Morton, A., Sharples, L., & Taylor, A. (2021, June 1). Reviewing the evidence: heuristics and biases. Developing a Reference Protocol for Structured Expert Elicitation in Health-care Decision-making: A Mixed-methods Study — NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK571047/

9 Deng, N., Hodroj, B., Latham, A. J., Lee-Tory, J., & Miller, K. (2024). Is present-bias a distinctive psychological kind? Inquiry, 1–27. https://doi.org/10.1080/0020174x.2024.2321614

10 See https://poverty-action.org/blog/mobile-izing-savings-with-defaults-afghanistan

- **Empathy gaps** occur when people have difficulty anticipating how they —or someone else— will think, feel, or act in a different emotional state or situation.[11] One way to reduce these gaps is by guiding individuals through a deliberate process of imagining another person's experience or considering how they themselves might respond in a different context

  For example, in Turkey, a year-long school-based curriculum introduced perspective-taking activities in elementary schools where 18% of students were refugees. The programme encouraged students to adopt the perspectives of out-group peers[12], fostering both emotional and cognitive empathy. As a result, schools achieved significant reductions in peer violence, social exclusion, and ethnic segregation, alongside gains in trust, altruism, and other pro-social behaviours. This example shows how early interventions can help bridge empathy gaps and promote long-term social cohesion in diverse communities.[13]

- **Fundamental attribution error** refers to the tendency to overemphasise personal traits (such as someone's character, intentions, or abilities) and underemphasise situational factors (such as context, environment, or circumstances) when explaining other people's behaviour. Conversely, when people explain their own behaviour, they often attribute it to the situation, rather than to their character or disposition. This asymmetry — judging others by who they are, while judging ourselves by the circumstances we face — shapes many interactions.

  Consider another hypothetical health situation in which a mother brings a very sick child to the clinic. The health care worker might think, "Mothers bring their children to the clinic too late because they don't care. I must scold them to bring their children in at the first sign of illness!" While this frustration is understandable, this view overlooks situational barriers such as long travel distances, high costs, or competing responsibilities. Here, fundamental attribution error leads to blaming the person rather than understanding their context.

- **Confirmation bias** is the tendency to selectively notice, value, pay attention to, and believe information that supports our existing beliefs. While this focus helps us to filter what feels most relevant, it can also reinforce inaccurate views and harmful behaviours.

  Researchers have long studied a particular type of confirmation bias among teachers: how prior expectations about student performance shapes their perception and subsequent treatment of students, potentially creating a self-fulfilling prophecy of low (or high) achievement. While much of this work has focused on disadvantaged groups in the U.S.,[14] recent studies are examining the issue in diverse contexts worldwide.[15]

The examples in Lesson 2 of the training materials provide more examples and explanations of common biases, drawing on studies of how they shape decision-making.[16],[17]

While heuristics and biases can resemble habits, they are different. Biases are systematic shortcuts in judgment, while habits are behaviours triggered automatically by environmental cues, often occurring outside conscious

11 For one of the seminal papers on this topic, see Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. Health Psychology, 24(4, Suppl), S49–S56.

12 According to the American Psychological Association's Dictionary of Psychology, outgroup refers to "any group to which one does not belong or with which one does not identify".

13 Chatterjee, Shreya; Gassier, Marine; Myint, Nikolas. Leveraging Social Cohesion for Development Outcomes (English). Policy Research working paper ; no. WPS 10417 Washington, D.C. : World Bank Group.

14 See Jussim and Harber (2005) for a summary. Jussim, L., & Harber, K. D. (2005). Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies. Personality and Social Psychology Review, 9(2), 131–155.

15 E.g., Zadja (2021). Zajda, J. (2021). Discrimination and Self-Fulfilling Prophecy in Schools Globally. In J. Zajda (Ed.), Globalisation and Education Reforms: Creating Effective Learning Environments (pp. 51–70). Springer International Publishing.

16 Casigliani, V., Menicagli, D., Fornili, M., Lippi, V., Chinelli, A., Stacchini, L., Arzilli, G., Scardina, G., Baglietto, L., Lopalco, P., & Tavoschi, L. (2022). Vaccine hesitancy and cognitive biases: Evidence for tailored communication with parents. Vaccine X, 11, 100191. https://doi.org/10.1016/j.jvacx.2022.100191

17 Azarpanah, H., Farhadloo, M., Vahidov, R., & Pilote, L. (2021). Vaccine hesitancy: evidence from an adverse events following immunization database, and the role of cognitive biases. BMC Public Health, 21(1). https://doi.org/10.1186/s12889-021-11745-1

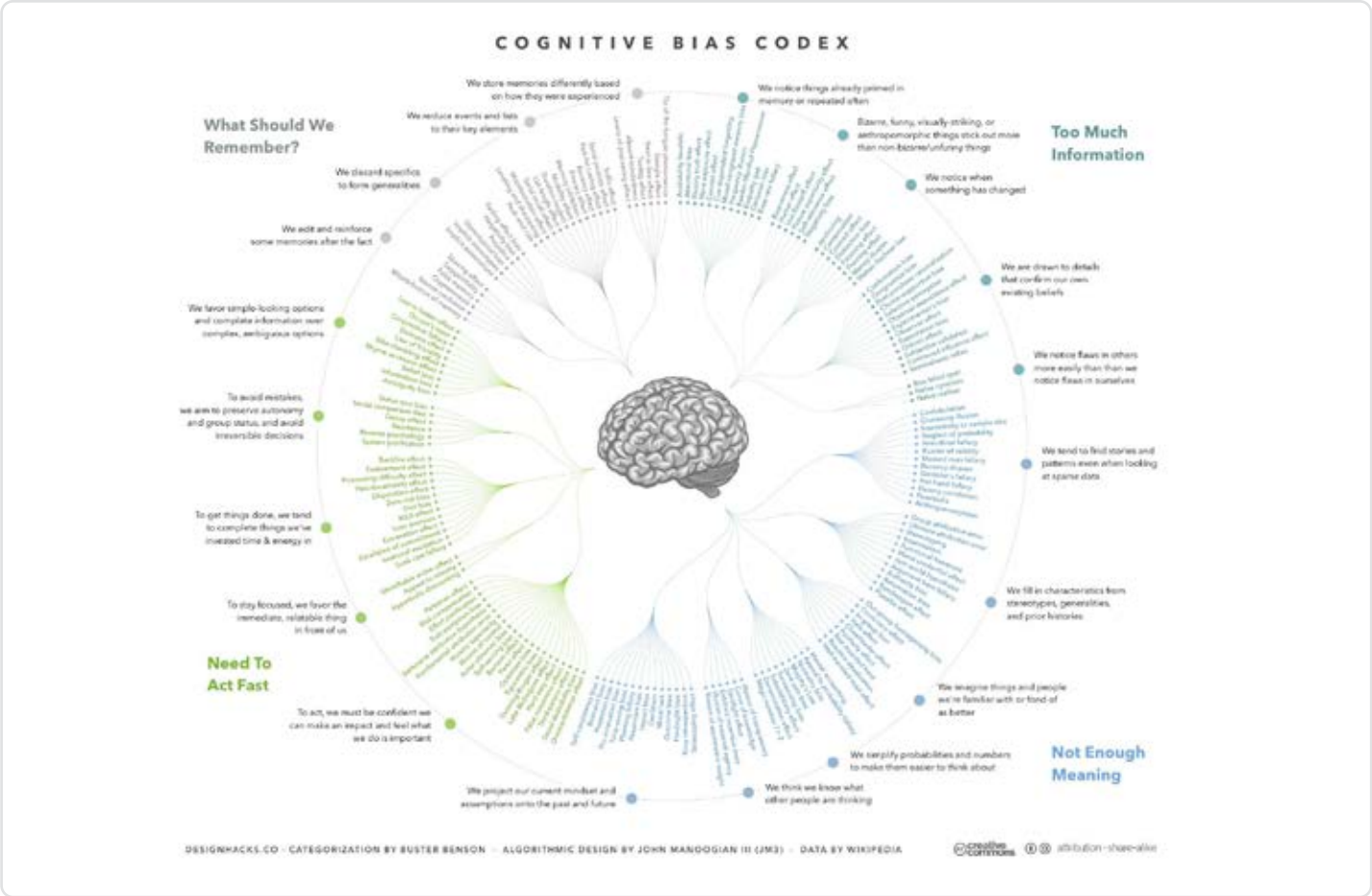


Figure 4: Cognitive Bias Codex (Source: School of Government, UNC)

control. Although habits are not usually listed as biases or heuristics, they can be seen as one of the most powerful shortcuts, helping us to conserve mental effort.

Habits shape daily life, including the lives of families and children. In childhood nutrition, for example, the foods purchased and prepared at home often follow ingrained routines — programmes that ignore these patterns are unlikely to succeed.[18] Research in Kenya likewise shows that handwashing, when established as a non-conscious habit, is a key factor in whether parents transmit childhood infectious diseases.[19]

18 See for example Ji and Wood (2007). Ji, M. F., & Wood, W. (2007). Purchase and Consumption Habits: Not Necessarily What You Intend. Journal of Consumer Psychology, 17(4), 261–276. https://doi.org/10.1016/S1057-7408(07)70037-2

19 See Aunger et al. (2010). Aunger, R., Schmidt, W.-P., Ranpura, A., Coombes, Y., Maina, P. M., Matiko, C. N., & Curtis, V. (2010). Three kinds of psychological determinants for hand-washing behaviour in Kenya. Social Science & Medicine, 70(3), 383–391. https://doi.org/10.1016/j.socscimed.2009.10.038

When do people rely on shortcuts?

Mental shortcuts are central to how the mind works, reducing the effort of navigating countless daily decisions. People are most likely to depend on them in challenging situations, like the following:

- **Complex decisions** with too much information to process, many options to weigh, or high uncertainty about future outcomes.
- **Distractions and hassles** that interfere with decision making and swift action, including too many steps, a long delay between decision and action.
- **Strong social pressure** to take a particular action or decision.
- **Repetition of a situation**, where a similar choice has been confronted many times before, without facing any negative consequences for prior decisions.
- **Low importance**, when the choice simply isn't a priority or is far less important than other more pressing matters.

Shortcuts aren't inherently bad — they are clever strategies that often serve us. However, they can be misleading when people face choices that are more difficult, less familiar, or perceive that they are not supported by other members of the community or their social group. These are exactly the situations where behavioural science can add the most value.

What implications does this have for changing behaviour?

Figure 5: Intention-Action Gap. Busara 2025.

The first implication is the gap between intentions and actions. People may want to do or accomplish something of importance, but fail to follow through. For example, a caregiver may know the value of clinic testing for a chronic disease, yet present bias shifts their attention to immediate tasks like preparing meals or getting a child to school.

Second, behavioural science affirms that the human mind isn't defective, and people who appear to act irrationally aren't foolish. As researchers and practitioners, we sometimes fall into the trap of blaming others for not making the "best" choice, partly due to our own fundamental attribution error. We overlook that everyone's minds are limited, including our own, and that the shortcuts we use are generally clever and useful, even if imperfect.

Third, behavioural science shows it is often more effective to change the context than to change people's minds. By examining the system around a problem, it's possible to redesign environments so better choices are easier to make. Analysing context and developing practical solutions is the essence of applied behavioural science, and the focus of this guide.

Taken together, these insights can help us to design interventions that address the real roots of why people do what they do.

3. UNICEF's Approach: DEPTHS

About DEPTHS

UNICEF has developed DEPTHS, a systematic process to help guide practitioners apply behavioural science in practice. DEPTHS stands for Define, Explore and Diagnose, Prototype Designs, Test Hypotheses, and Scale. At a high level, it mirrors the common blueprint used by most behavioural science organizations and researchers, bringing scientific rigour to real-world challenges.[20]

Most of the commonly used approaches in applied behavioural science were developed years ago, before the current level of integration across disciplines. Today, best practices in behavioural science draw from multidisciplinary approaches, such as social psychology, cognitive science, experimental economics, anthropology, design, systems analysis, among many others.

This guide draws on existing resources across disciplines, while highlighting the aspects of applied behavioural science that are most relevant for working with and for children. In particular, it emphasizes where applied behavioural science offers something distinct, such as mapping micro-behaviours, or identifying biases that influence survey responses. In doing so, it builds on established approaches in SBC, Human-Centred Design (HCD), and related fields, while adding new tools, mindsets, and evidence to the shared toolbox for ethical behavioural impact.

20 Wendel, Khan and Artavia-Mora (2023). For example, BIT uses a process called TESTS, Busara uses AUDAS, and Ideas42 uses IDEAS. Each has five steps, and each step has similar objectives as DEPTHS.

A brief look at each step

The remaining chapters in this field guide provide a detailed look at each step of DEPTHS, covering their respective purposes, practical tools and templates, illustrative case studies, and suggestions for further reading. The next chapters cover the full DEPTHS process, as outlined below:

D **Define:** Ensuring that researchers clearly identify the specific problem and the particular behaviours they aim to address with an intervention.

E **Explore and Diagnose:** Using research and data to understand the sociocultural, environmental, and psychological factors at play.

P **Prototype Designs:** Using human-centred design, prototyping, and participatory methods to co-create and pre-test tentative solutions directly with the people who will use them.

TH **Test Hypotheses:** Piloting interventions in real-world conditions to see what works, adapt designs, and generate evidence before scaling.

S **Scale:** Expanding effective interventions to new contexts and populations and embedding them into policies and systems for long-term impact.

To bring the process to life, this guide uses a case study from Lebanon as a thumbnail overview of how the DEPTHS process works in practice. After the Syrian crisis, many caregivers began but did not complete their children's immunization schedules. To address this, parents received leaflets asking them to commit to completing vaccinations and providing practical details on when and where services were available. This low-cost, behaviourally informed intervention was associated with a 6.8% increase in vaccine completion.

This case study is carried across the DEPTHS chapters to illustrate how the same challenge can be approached step by step: from defining the problem, to exploring behavioural drivers, to prototyping, testing, and scaling solutions.

Define

The journey of applied behavioural science begins here.

Before jumping into conducting research or designing solutions, it's essential to step back and define what success looks like.

This involves answering questions such as:

- What is the problem we're solving?
- What is the outcome we want to achieve?
- What is the specific behaviour that we want to see change?
- What is the system we are operating within?
- And finally, where are we best placed to make an impact?

Why *Define*?

Too often, teams jump straight to designing or implementing solutions without a shared understanding of the underlying issue. This could be due to resource constraints, tight deadlines, or lack of capacity, among other reasons. However, without a clear and shared problem definition, efforts risk being inefficient.

When faced with complex challenges in humanitarian and development work, it can be tempting to exclusively prioritize broad and systemic reforms. Change can also start with identifying and influencing specific behaviours. Breaking down overwhelming problems into concrete behavioural questions can reveal actionable solutions and deepen understanding of what drives change.

How can we *Define* the challenge?

In this first phase of the DEPTHS methodology, there are five steps to: build a shared understanding of the problem; identify a specific behaviour to change; and pinpoint leverage points to focus on. While most steps are essential to ensure a high-quality behavioural science project, others are optional. Optional steps will be noted.

Each step includes specific tools, guidance on why these tools matter, and how to use them. A **case study on increasing childhood vaccination uptake in Lebanon** will exemplify how the tools can be applied in practice.

Steps in the *Define* phase

1 Define the problem and outcomes

Specify the problem, outcome(s) of interest, and which stakeholders and population groups to engage.

TOOLS:

- Problem Definition and Outcome Statement
- Stakeholder and Audience Map

2 Identify potential target behaviours

With the problem and outcomes defined, identify the constituent behaviours. Select one that has the highest potential impact and feasibility of being changed

TOOLS:

- Behaviour Tree
- Impact-Feasibility Prioritization Matrix

3 Map the system (OPTIONAL)

Create a graphic representation of the relationships between different factors in the system. Include both drivers and barriers to success in addressing the problem and achieving the desired outcome.

TOOLS:

- System Map

4 Find leverage (OPTIONAL)

Use the system map to identify leverage points: aspects of a system that, if changed, would have a high impact on the desired outcomes.

TOOLS:

- Leverage Point Analysis

5 Document the project and scope

All of the outputs of the *Define* phase will culminate in a one-page summary: the project canvas. This document outlines the intended outcome, leverage points, potential actors, and actions that will drive change, along with roles and responsibilities.

TOOLS:

- Project Canvas

The Define phase requires a committed team and adequate resources. Depending on the project's goals, it may also help to involve specialists in areas such as experts from the programme itself, social and behaviour change, human-centred design, or systems thinking.

Common pitfalls

There are common traps that can derail behaviourally-informed work. Throughout the *Define* phase, keep these possible issues in mind:

- **Rushing to solutions.** Jumping straight into solutions or interventions without considering all of the factors can lead to ineffective or misaligned programmes. For example, designing an app to help caregivers track their children's vaccination schedule might seem promising, but without first understanding how caregivers access information, what motivates follow-through, or whether they trust digital tools, the solution may fall flat. The best solutions often emerge after framing the problem and learning more about the context.
- **Failing to engage communities.** Defining the problem and identifying priorities without community engagement and participation can result in a mismatch between the project's focus and what matters to those who are directly affected. For example, an immunization programme might prioritize building more fixed-site clinics to improve access, assuming proximity is the main barrier. However, if communities aren't consulted, the programme might overlook the reality that caregivers often fear mistreatment by health workers, or that household duties make clinic hours inaccessible. Community engagement isn't a checkbox; instead, it grounds applied behavioural science in real-world experience.
- **Blaming the individual.** Avoid framing problems as what people are doing "wrong." Focus instead on forces that shape behaviour: structural, cultural, or environmental barriers. For example, a campaign encouraging caregivers to vaccinate on time might start with 'negligence', yet caregivers may face long travel distances, missed wages, or waiting long hours at the clinic. Environmental factors like transport, socio-cultural drivers like family community and expectations, and psychological drivers such as trust in the health system all influence what people do or do not do.
- **Solving symptoms without focus.** Goals like 'increase contraception uptake' are too broad. Instead, focus on specific behavioural objectives (e.g., "increase consistent contraceptive use among adolescent girls in peri-urban areas"). This makes it easier to identify root causes, design tailored interventions, and measure progress.
- **Treating assumptions as facts.** Flag assumptions and treat them as hypotheses to test, not conclusions to act on. For example, introducing a financial penalty for parents who arrive late to pick up their children may seem like a logical deterrent. Yet this can backfire as parents may view the fine as a fee they're willing to pay, rather than a behaviour to avoid.
- **Ignoring power dynamics.** There is always a risk of reinforcing existing hierarchies or overlooking who actually holds influence over decisions. If formal and informal decision-makers are left out, solutions risk being ineffective. In nutrition programmes, focusing messaging solely on mothers may ignore the authority of husbands, grandmothers, or local leaders in shaping food choices.
- **Skipping documentation and alignment.** Without clearly documenting the behavioural focus, population of interest, scope, and open questions, teams risk misalignment, duplication of work and a lack of focus. D key decisions clearly to keep the team aligned, and provide a shared reference as the project evolves.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

Childhood vaccination is one of the most cost-effective public health interventions, yet many children, especially in low- and middle-income countries (LMICs), remain un- or under-vaccinated. In Lebanon, home to the world's highest per-capita refugee population, challenges to vaccination uptake are compounded by poverty, displacement and strained health systems. Although national immunization coverage once neared 90 per cent, outbreaks of measles and mumps in 2013 and 2015 revealed growing pockets of under-vaccination.

In response, the Ministry of Public Health and UNICEF launched the Accelerated Immunization Activities (AIA) programme to reach vulnerable children through both health centres and community outreach. Yet even with free services and education campaigns, under-vaccination was still high.

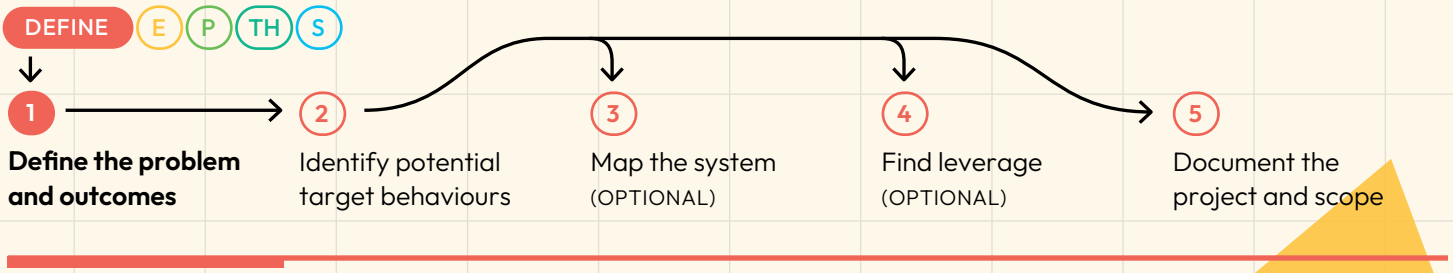
A multidisciplinary team from UNICEF, Nudge Lebanon, and the Ministry of Public Health (MoPH) set out to explore these challenges from a behavioural science perspective to define the issue Rather than focusing solely on structural or access barriers, they focused on caregiver decision-making and identified barriers such as long wait times, socio-cultural constraints, and trust in the health system.

The team grounded their work in real-world insights and set the stage for testing interventions by articulating the problem through the lens of caregiver decision-making.

Throughout this guide and starting with the Define phase, the Lebanon research team's journey will serve as a real-world example of how each phase of the DEPTHS process can help teams to effectively apply behavioural insights.

Note: *While this team did not explicitly use all the tools in the DEPTHS field guide, real project data has been used to illustrate and align their design process.*

© UNICEF2021/Fouad-Choufany/Lebanon



STEP 1:

Define the problem and outcomes

In this step:

This step guides teams to define the problem and write a SMART (Specific, Measurable, Achievable, Relevant, and Time-bound) Outcome Statement. The outcome should be ambitious yet realistic, grounded in what matters to the people UNICEF serves, and feasible for the project team. A strong outcome statement defines the:

- Specific change to be achieved
- Target population for behaviour change
- Magnitude of the intended change
- Timeframe for the intended change

Associated tools:

- Outcome Statement
- Stakeholder and Audience Map

Why it matters:

A well-defined problem is more than just a topic or general concern, such as 'low immunization rates' – it's a specific, observable gap between how things are and the desired behaviour. Defining the problem means asking: Who is affected? What is going wrong? Why does it matter? What is the cost of inaction?

That's the role of an outcome or objective: a clear, tangible result that the project will work toward, within the time, scope, and resources available. This ensures outcomes remain grounded in measurable results for individuals, not just institutional priorities. A strong outcome statement clearly articulates the change the project is working toward. It prevents premature 'solutioning' and serves as a guiding anchor for every decision that follows.

This is also a good moment to check whether behavioural science is the right fit for your problem: Applied behavioural science helps most when people have capability, opportunity, and motivation to act, but still face barriers to follow-through.

How to do it:

1. Work with experts to pinpoint opportunities

Start by consulting colleagues or sector specialists. In most cases, persistent challenges have already been identified, whether they're a bottleneck in service delivery, a gap in uptake, or an issue that hasn't responded to existing approaches. Sector specialists and expert colleagues can help pinpoint where efforts have stalled and identify areas where a behavioural sciences approach could add the most value.

In the case of increasing childhood vaccination uptake in Lebanon, the problem was already being addressed by the Ministry of Health through the accelerated immunization activities (AIA) programme. However, specialists recognized that this effort alone was not enough to increase immunization rates.

Be wary of becoming too narrow in focus. Familiarity with a subject can lead to assumptions or blind spots about the wider system. While experts provide depth, their perspectives may be biased. That's why it is essential to validate assumptions with local stakeholders in the next steps, such as Step 2 later in this chapter: Identify potential target behaviours.

2. Articulate the problem

After gathering insight from both experts and local stakeholders, it's time to use the left side of the Problem Definition and Outcome Statement worksheet to clearly articulate the problem the project will address.

a. Questions 1–3 describe the problem: What's happening, where is it happening, and why does it matter? Use insights from both expert input and community perspectives.

b. Use data where possible to illustrate the size of the problem (e.g. prevalence, coverage rates, or service gaps)and to help anchor the issue in measurable terms.

c. Pause and reflect: Is this the right problem to focus on? Whose perspectives are being represented? What assumptions might we be making?

d. Finally, assess whether the problem is a good fit for a behavioural science approach by asking: Could changing decisions, behaviours, or the context in which they occur, lead to meaningful impact or desired outcomes?

In formulating the problem statement, make sure to:

- **Be specific about *who* is affected and *where*.** Avoid generic terms like 'community member' or 'young people' unless clearly defined. Instead, specify populations of interest — for example, "adolescent girls aged 10–19 living in rural areas".
- **Focus on observable gaps.** What is happening that shouldn't be, or what should be happening that isn't?
- **Describe an actual problem, not just the absence of an activity.** For example, 'low uptake of vaccines' is a problem. In contrast, 'lack of health worker training' or 'lack of incentives' are potential causes or contributing factors, but not problems themselves.
- **If possible, refer to existing data or evidence that supports the problem description.** This will lay the groundwork for a stronger diagnosis later.

Problem Definition

What is the issue we want to solve?

1 **What is the problem we're seeing?** *Briefly describe the issue and its broader context. What's happening, where, and why does it matter? Include data, such as prevalence, coverage, or service gaps, to help anchor the issue in measurable terms.*

2 **How do we know it is the right problem?** *Whose perspective is reflected? What assumptions are we making?*

3 **Can this problem be addressed by using a behavioural science approach?** *Could changing decisions, behaviours, or context make a meaningful difference?*

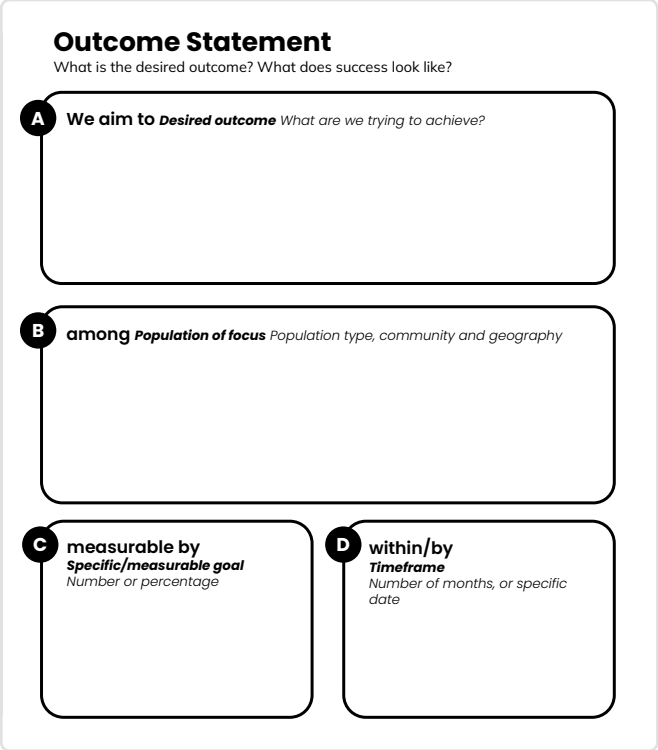
3. Create the outcome statement

Move to the right-hand side of the Problem Definition and Outcome Statement to ask: What specifically are we trying to change by addressing this problem? The outcome statement should directly respond to the problem and reflect the measurable improvement if the project is successful.

Write a SMART goal:

- **Specific:** clearly defined
- **Measurable:** linked to observable changes
- **Achievable:** realistic for the project team and context
- **Relevant:** aligned with the scope and priorities
- **Time-Bound:** with a clear timeframe for change

Specifically, the following areas should be filled in: the desired outcome (the change or action the project is aiming to achieve), the population of focus (the population type, community, and geography), the specific and measurable goal (ideally expressed in numbers or percentages), and the timeframe (such as the number of months or a specific end date). The table on the following page provides examples of weak and strong outcome statements.



While crafting the outcome statement, keep these points in mind:

- **'Awareness' is not an outcome.** Focus on measurable, observable changes in the real world – metrics that can be tracked over time. For example, caregivers may already be aware of the importance of vaccines; this awareness doesn't guarantee that caregivers will follow through on appointments or complete their child's vaccination schedule on time.
- **The outcome should not yet include a specific behaviour.** Concentrate on defining the high-level result the project seeks to achieve. That said, if a potential target behaviour emerges during this part of the process, make a note of it. It may be useful to revisit, upon reaching 'Step 2: Identify potential target behaviours'. For example, if the goal is to increase timely childhood vaccination, the aim might be to support caregivers in making a plan for vaccination appointments, or health workers in following up with families more consistently. However, these behavioural aspects come later. Instead, the focus should be on articulating the broader outcome that the project is working towards.
- **Outcomes should describe the desired change, not the proposed solution to achieve it.** Before jumping ahead to specific interventions, fully explore the problem's context through the behavioural science process. Let the outcome guide thinking, not pre-empt it.
- **Outcomes should account for existing gender, socio-cultural, and power dynamics.** Pause to reflect: Are there differences in roles, norms or power dynamics that affect how different groups experience the problem we're addressing? Consider framing an outcome that not only tackles the core issue, but also helps reduce underlying inequities, whether based on gender, age, ethnicity, disability, or other forms of marginalization.

TABLE 1. WEAK VS STRONG OUTCOMES (SMART)

WEAK EXAMPLE	WHY IT'S WEAK	STRONG EXAMPLE
"Enhance school retention of adolescent girls."	It's not clear what "school retention" means in practice: are we talking about primary, secondary, or both? There's no specific location, age group, or timeframe, and "enhance" is too general to track progress.	Increase secondary school attendance and reduce dropout rates among adolescent girls in rural areas of [area name], measurable by a 20% reduction in dropout by grade 9, within 2 academic years.
"Increase vaccination rates in target areas."	Clear intent, but too broad: no baseline, no target %, no timeframe, no defined population (Which vaccine? Which age group?).	Improve routine childhood vaccination coverage among children under 2 in low-coverage districts of [area name], measurable by a 10 percentage point increase in full immunisation rates, within 12 months.
"Promote good nutrition practices amongst children."	Activity-framed rather than outcome-focused. "Promote" implies process, not result. Also doesn't clarify who or what success looks like.	Improve early identification and treatment of malnutrition among children under 5 in drought-affected regions, measurable by a 25% increase in admissions to community-based management of acute malnutrition services in [area name], within 9 months.

4. Map stakeholders and target audiences

The **Stakeholder and Audience Map** worksheet helps to clarify which individuals are most interested in or concerned by the problem, along with their influence. This tool helps to map these groups of people and consider which type of engagement will ensure future research and intervention efforts are well targeted.

Use the matrix in the **Stakeholder and Audience Map** worksheet to sort individuals and groups based on how much power they hold to shape the issue, and how actively involved or interested they are. Mapping stakeholders this way can help inform research recruitment, partnerships, and engagement strategies. It can also aid in choosing the most appropriate engagement strategy.

Place stakeholders on the matrix according to:

- **Influence:** To what degree can this stakeholder support or hinder the problem?
- **Involvement:** To what extent is this stakeholder interested in or concerned about the problem?

TABLE 2. STAKEHOLDER MAP CRITERIA AND ENGAGEMENT STRATEGIES

INFLUENCE	INVOLVEMENT	ENGAGEMENT STRATEGY
High	High	**Manage:** These stakeholders are both influential and highly invested in the issue. Prioritize active collaboration and manage their involvement closely. **Case example:** The technical project team in charge of increasing immunization rates within the Ministry of Health of Lebanon.
High	Low	**Satisfy:** These stakeholders are influential but less engaged. Keep them informed and find ways to maintain their support without overburdening them. **Case example:** High level public officials within the Ministry of Health of Lebanon.
Low	High	**Inform:** These groups are very interested but may have limited influence. Keep them updated and draw on their lived experience or technical insight to inform the work. **Case example:** Caregivers of young children, including both Lebanese and Syrian refugee families.
Low	Low	**Monitor:** These stakeholders are currently less relevant but may become more engaged over time. Stay aware of shifts in interest or influence. **Case example:** Local NGOs working on general child health or maternal care, but not currently involved in the immunization programme.

Use the second half of the Stakeholder and Audience Map worksheet to brainstorm and list target audiences:

- **Primary audiences:** Individuals or groups who experience the problem firsthand, who directly influence the targeted outcome through their actions, or who will most likely benefit from a solution. Examples include caregivers of children who need to get vaccinated, parents of children with malnutrition, adolescent girls at risk of school dropout, or residents of informal settlements with limited access to clean water.
- **Secondary audiences:** These might be individuals or groups who indirectly shape the target outcome through decisions, habits, or roles. They influence the environment, decisions, or access of the primary audience. This may include gatekeepers (e.g., community elders, religious leaders, or health facility managers), influencers (e.g., peers, family members, or teachers), and decision-makers (heads of households, community leaders, or authorities). Though they may not always be directly affected by the issue, their roles shape how primary audiences experience it. Their influence may be supportive or detrimental to the outcome.

When mapping out populations of interest:

- **Avoid over-generalising.** Using terms like 'caregivers' or 'women' can mask key differences within a group. Consider factors such as age, gender, marital status, geography, and socioeconomic background when segmenting audiences.
- **Look beyond the most visible actors.** The loudest or most visible groups aren't always the most impacted. Be intentional about surfacing marginalized voices who may experience the issue differently.

Target Audience

The specific populations most connected to the problem at hand.

Primary audience

Who is directly affected by the problem, is experiencing the issue firsthand or stands to benefit most from a solution?

Secondary audience

Who influences the decisions, access, or environment of the primary audience? Who are the gatekeepers, influencers, or decision-makers shaping their experience?

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

This Problem Definition and Outcome Statement was not developed by the original project team. It is a recreated example based on real project data and context. Its purpose is to illustrate what a completed Problem Definition and Outcome Statement worksheet could look like, in practice.

The project team was concerned by emerging evidence of missed vaccinations among both host communities and refugee populations, an issue that became more urgent after a rise in cases of measles and mumps. Despite the availability of free vaccines through primary healthcare centres, surveys conducted by the MoPH revealed clusters of under-immunized children, especially among families facing socio-cultural or logistical barriers.

To better understand the problem, the team reviewed health administrative data and engaged in conversations with health workers, outreach staff, and caregivers. The project team developed a clear outcome statement focused on increasing on-time vaccination among children under five in communities targeted by Lebanon's Accelerated Immunization Activities (AIA) programme.

Stakeholder mapping led to decisions to maintain a close and collaborative relationship with the technical team at the MOPH, given their high level of influence and involvement. High-level public officials – while influential – were kept informed without being overwhelmed by frequent updates or requests, ensuring their support without overburdening them.

Finally, researchers began brainstorming and mapping the primary and secondary audiences of the project. The primary audience included caregivers of un- or under-vaccinated children. Secondary audiences included,health workers, outreach teams, and community leaders. Their roles as trusted messengers, gatekeepers, or facilitators made them critical allies in shaping supportive environments for vaccination.

DEFINE E P TH S

DEPTHS TOOLKIT

CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

Stakeholder Map

The people most concerned by the problem as well as how much influence they have on it.

Influence
To what degree can stakeholders support or hinder the problem?

High

SATISFY
High Influence, Low involvement

1. High level public officials within the Ministry of Health of Lebanon.
2. Representatives from international donor agencies funding health programs in Lebanon (e.g., Gavi, the Vaccine Alliance).

MANAGE
High influence, High involvement

The technical project team in charge of increasing immunization rates within the Ministry of Public Health of Lebanon.

Low — High

Involvement
How interested or concerned are stakeholders in the problem?

Local NGOs working on general child health or maternal care, but not currently involved in the immunization programme

Caregivers of young children, including both Lebanese and Syrian refugee families.

MONITOR
Low Influence, Low Interest

Low

INFORM
Low Influence, High Interest

Target Audience

The specific populations most connected to the problem at hand.

Primary audience

Caregivers of un- or under-vaccinated children: Primarily mothers, but also fathers, grandparents, or other household members responsible for child health.
Un- or under-vaccinated children (ages 0–16): Children identified as behind on Lebanon's routine immunization schedule.

Secondary audience

- **Outreach workers / AIA community health teams** who conduct home visits, provide education, referrals, and follow-ups.
- **Primary Health Center (PHC) staff** who deliver vaccinations and track immunization records
- **Community leaders or gatekeepers** (e.g., mukhtars, religious leaders, or local influencers) who shape social norms and trust
- **Older family members** (e.g., grandparents) who may influence caregiver decisions based on generational beliefs or past experiences
- **Peers and neighbors** who can affect perceived norms and influence decisions through informal conversations
- **Local NGOs** involved in health promotion or refugee support in the region)

STEP 2:

Identify potential target behaviours

In this step:

Two tools support this step: the Behaviour Tree (to map behaviours) and the Prioritization Matrix (to select the most promising):

1. **Behaviour Tree worksheet:**
   This is a visual map that links the desired outcome to:
   - **a.** Key audiences (e.g., parents, health workers, teachers, or religious leaders).
   - **b.** The behaviours — or lack thereof — that affect the outcome for each audience group.

   The goal is to show how different individuals and their actions — or inactions — contribute to the issue. The Tree highlights where to position the intervention, in order to have the greatest impact.

Associated tools:

- Behaviour Tree
- Prioritization Matrix

2. **Prioritization Matrix worksheet:**
   This tool helps to narrow focus by comparing behaviours from the Behaviour Tree against two criteria:
   - **a.** **Feasibility**: How realistic is it to influence or change this behaviour?
   - **b.** **Impact**: How much would changing this behaviour contribute to the desired outcome?

   Many behaviours are important, but this matrix helps to prioritize one or two that are achievable and high impact.

Why it matters:

Behavioural science is fundamentally about understanding and influencing what people do. Without a clearly defined target behaviour, it is not possible to meaningfully investigate what drives or hinders change, design interventions that are likely to have impact, or evaluate whether efforts are achieving the intended effect.

It's easy to confuse behaviours with things like attitudes, emotions, or intentions. While these factors can influence behaviour, they are not behaviours themselves. A behaviour is a specific, observable action that a person takes. Not what they think, feel, or believe — rather, what they do. The table below shows how to reframe non-behaviours into true behaviours:

TABLE 3. BEHAVIOURS VS NON-BEHAVIOURS

NOT A BEHAVIOUR	WHY IT ISN'T A BEHAVIOUR	REFRAMED AS A BEHAVIOUR
Making caregivers understand the importance of vaccination	**Knowledge vs. Action:** While important, understanding is not something one can observe or reliably measure directly.	Caregivers take their child to a primary health centre within seven days of receiving a referral.
Increasing trust in vaccines among parents	**Belief vs. Action:** Trust influences action, but it can't be measured without interpretation.	Caregivers accept the vaccination without expressing hesitation when offered by the outreach worker.
Getting fathers more involved in vaccine decisions	**Vague vs. Specific:** 'Involvement' doesn't specify what action or decision the father is actually taking.	Fathers accompany the caregiver to the clinic or participate in outreach visits during immunization week.
Encouraging outreach workers to engage better with families	**Vague vs. Specific** 'Engage better' is ambiguous and subjective; it could mean many different things depending on the context.	Outreach workers explain the vaccination calendar and complete the referral form during every household visit.
Motivating caregivers to value routine immunization	**Intention vs. Behaviour:** Motivation is an internal driver; it's useful, but not directly observable or sufficient for behaviour change.	Caregivers mark the next vaccine date on the provided calendar and post it in a visible area at home.

How to do it:

1. Map behaviours

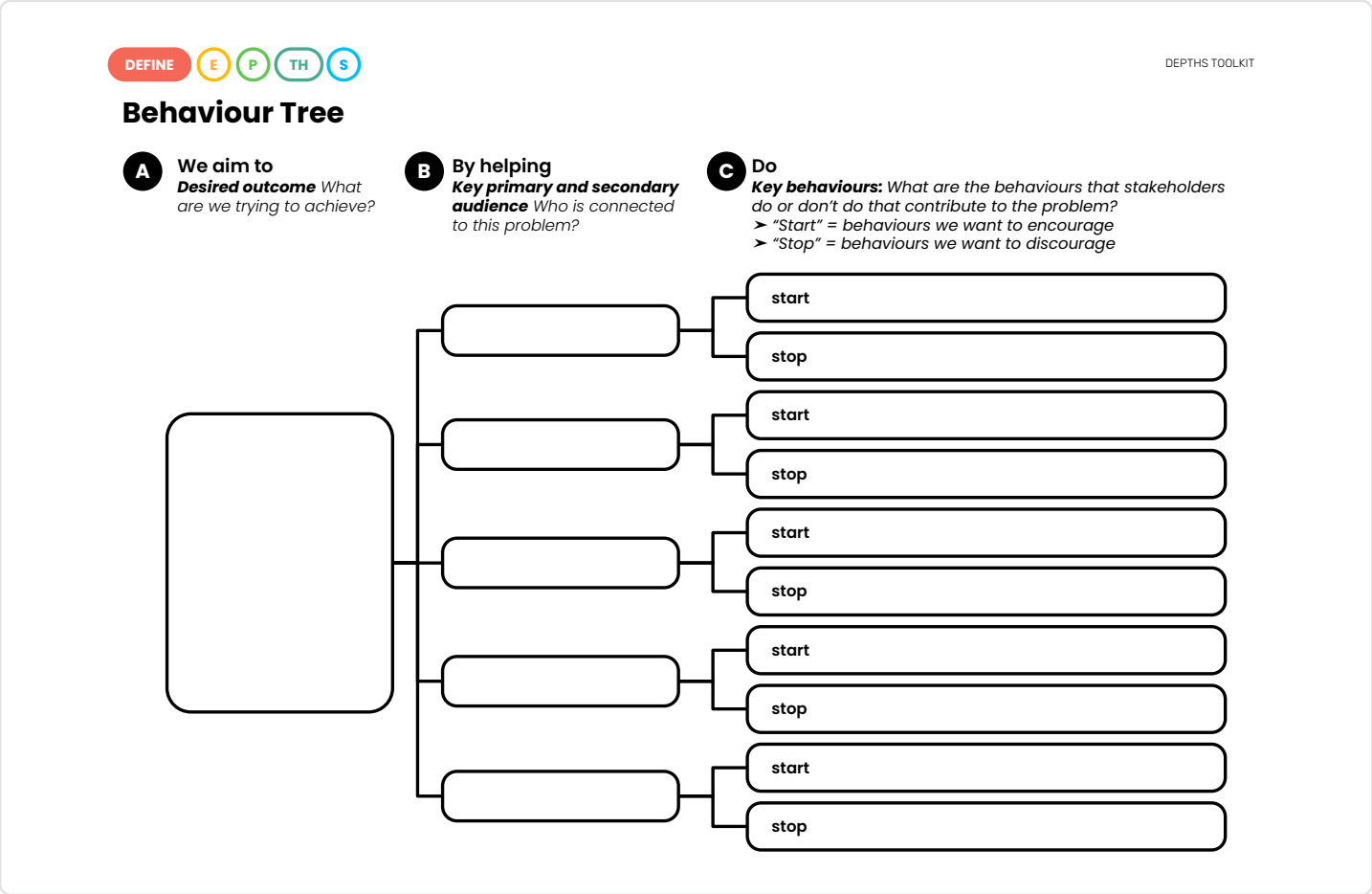
Use the **Behaviour Tree** worksheet to map out the behaviours that shape the target outcome:

a. Add the desired outcome in section A of the worksheet.

b. List the primary and secondary audiences identified in 'Step 1: Define the problem and outcomes' in section B.

c. For every audience listed in the Behaviour Tree, identify specific behaviours that influence the outcome in section C. This step is about surfacing both what's visible and what's missing in the behavioural landscape. Ask: What actions are each group currently taking that are helping — or hindering — progress? Are there helpful or positive behaviours these groups could be adopting, in order to support the project team to achieve its goals? Sort behaviours into two categories:

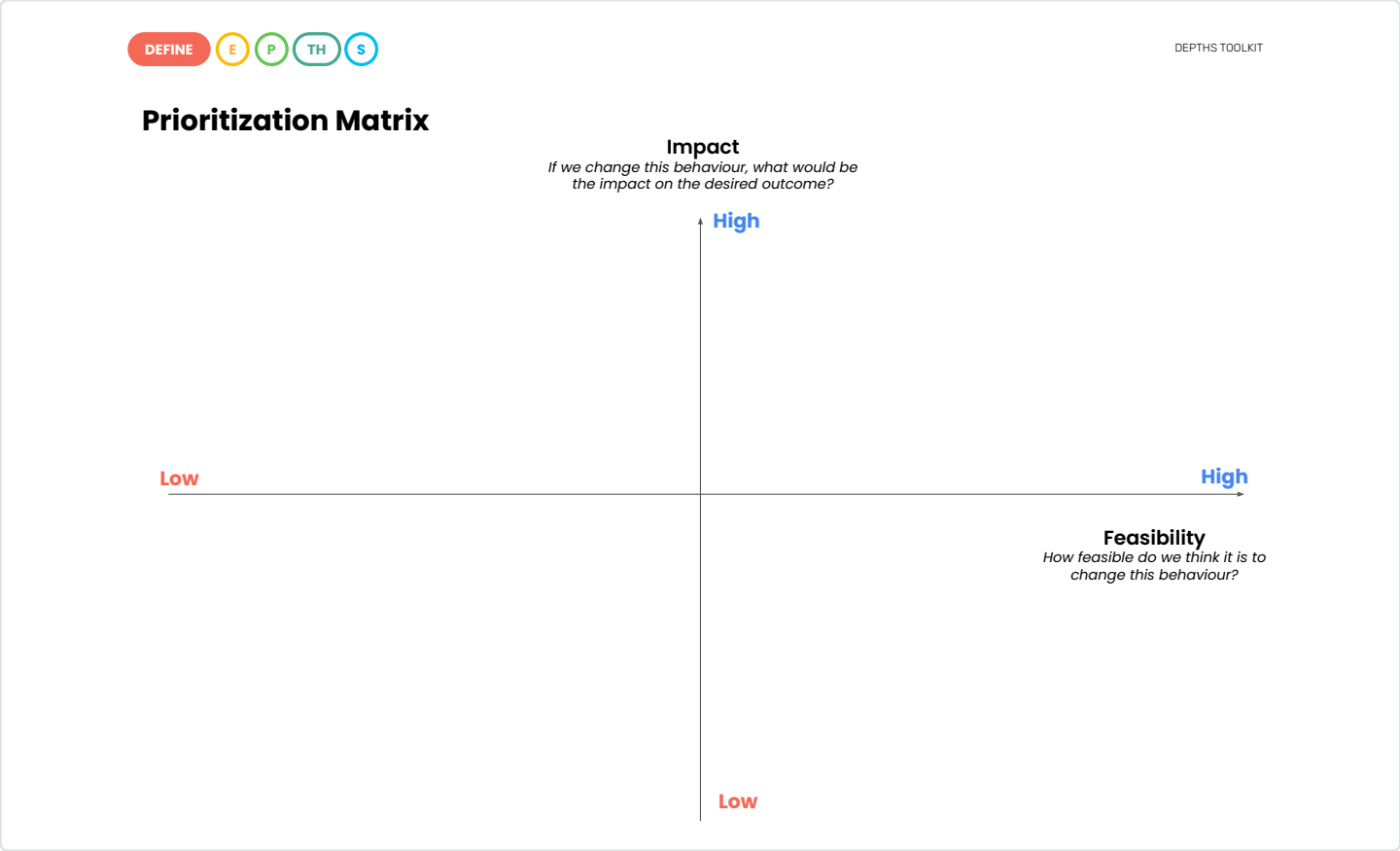
- **Start:** These boxes on the worksheet should include behaviours to encourage, which support the desired outcomes. These include both occurring behaviours that should continue, or new behaviours to take that further support the outcomes.
- **Stop:** These boxes should include behaviours to discourage, which are hindering the desired outcome.



To get the most out of the Behaviour Tree:

- **Start with a clear outcome.** The tree should always grow from a well-defined goal: desired outcome from Step 1.
- **If necessary, complement the initial list of audiences linked to the outcome.** The initial list of primary and secondary audiences from Step 1 should be comprehensive. This can be complemented by thinking broadly about people linked to the outcome.
- **Co-create whenever possible.** Local stakeholders and other team members bring lived experience that helps surface overlooked groups or behaviours, and prevents tunnel vision. There is no such thing as perfect information, but collaboration leads to a richer and more grounded understanding of the problem.
- **Focus on observable behaviours.** Remember, behaviours are things people *do*. If it can't be seen or measured, it's probably not a behaviour. Focus on what the person actually does, not what they think or feel.
- **Map what's there *and* what's missing.** Be sensitive to mapping both what people are doing and what they're not doing. This includes visible actions, like attending a health clinic or walking a child to school, as well as the absence of expected behaviours. It's often just as important to notice what's not happening: for example, caregivers not registering births, adolescents not attending counselling sessions, or men not accompanying partners to health visits.
- **Colour-code the Behaviour Tree.** Use a different colour for each stakeholder group and their specific behaviours. This will make prioritizing behaviours easier.

2. Prioritize behaviours



The **Prioritization Matrix** worksheet helps to sort through all of the identified behaviours and determine which ones are most worth acting on.

a. To start, take each of the behaviours from the **Behaviour Tree** worksheet and plot them on the matrix. Consider two factors:

- **Impact:** If this behaviour changed, how much would it improve the outcome?
- **Feasibility:** How realistic is it to change this behaviour, given the time, resources and context?

It might be helpful to think of a simple scale, like 'high', 'medium', or 'low', but expect to move things around. The matrix isn't static; it's a tool to organize your evolving thinking.

b. After placing behaviours on the matrix, select one to two to focus on for the next phase. This is usually a high-impact, highly feasible behaviour in the top-right quadrant, but it doesn't have to be. Real-world priorities like funding, partnerships, or existing initiatives may influence the prioritization of specific behaviours, and that's okay. The matrix is a guide, not a rule.

c. Finally, while prioritizing behaviours, try to avoid decision paralysis. Choosing a focus behaviour doesn't have to be perfect — it just needs to be thoughtful and grounded in what is known so far. Many teams refine their focus after doing deeper research in the next *Explore and Diagnose* phase.

3. Consult local stakeholders to validate assumptions

It's essential to pause and check any assumptions. True insight pairs technical expertise with lived experience. Too often, research agendas and funding priorities from the Global North set priorities on behalf of others. This can misalign efforts and overlook what matters most to communities.

Engaging local stakeholders, especially those closest to the issue, is one of the most effective ways to challenge biases. Estimating impact and feasibility is subjective, and the project team's perspective is just one piece of the puzzle. Involving colleagues, local partners, or people with different expertise can reveal new insights, challenge assumptions, and surface blind spots. Before moving to next steps, make sure to:

a. Discuss prioritized behaviours with the people most affected. Ask:

- Does this behaviour matter from their perspective?
- Are there parts of the problem that were misunderstood or overlooked?
- Are there other behaviours that feel more pressing or achievable?
- What would make this behaviour harder or easier to change?

b. Draw on participatory approaches. Validation doesn't have to be formal or extractive. It can range from informal conversations with community leaders to participatory techniques drawn from human-centred design[1].

A note on what comes next: This analysis may evolve if Steps 3 and 4 in the Define phase (system mapping and leverage points) are completed. These advanced steps explore root causes and influencing factors, and may refine the initial impact-feasibility assessment. Alternatively, teams can move directly to Step 5 and continue with the selected behaviour as the foundation for next phases.

1 "Human-centered design is a problem-solving technique that puts real people at the centre of the development process, enabling you to create products and services that resonate and are tailored to your audience's needs. The goal is to keep users' wants, pain points, and preferences front of mind during every phase of the process. In turn, you'll build more intuitive, accessible products that are likely to turn a higher profit because your customers have already vetted the solution and feel more invested in using it." — Landry, 2020, Harvard Business Review.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The Behaviour Tree and Prioritization Matrix were not developed by the original project team. They are recreated examples based on real project data and context

Caregivers in Lebanon sometimes postponed or skipped appointments due to uncertainty, low perceived urgency, or confusion about the vaccine schedule. The team recognized this as a behaviour **to stop** or discourage, and contrasted it with a behaviour **to start**: caregivers proactively keeping track of their children's vaccine schedule and ensuring timely completion of all required doses.

Outreach workers were also found to occasionally dismiss caregiver concerns or skip the repetition of key messaging during household visits — actions **to stop** that could erode trust or leave key information unclear. More constructive alternatives included **starting** consistent practices such as explaining which vaccines were missing, repeating important information, and engaging caregivers in a respectful, supportive way.

By sorting these real-world behaviours into 'start' and 'stop' categories, the team gained a more grounded view of the human actions driving uptake, and highlighted which behaviours could have the most influence on outcomes.

Behavior Tree

CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

A We aim to
Desired outcome *What are we trying to achieve?*

B By helping
Key primary and secondary audience *Who is connected to this problem?*

C Do
Key behaviours: *What are the behaviours that stakeholders do or don't do that contribute to the problem?*
➤ *"Start" = behaviours we want to encourage*
➤ *"Stop" = behaviours we want to discourage*

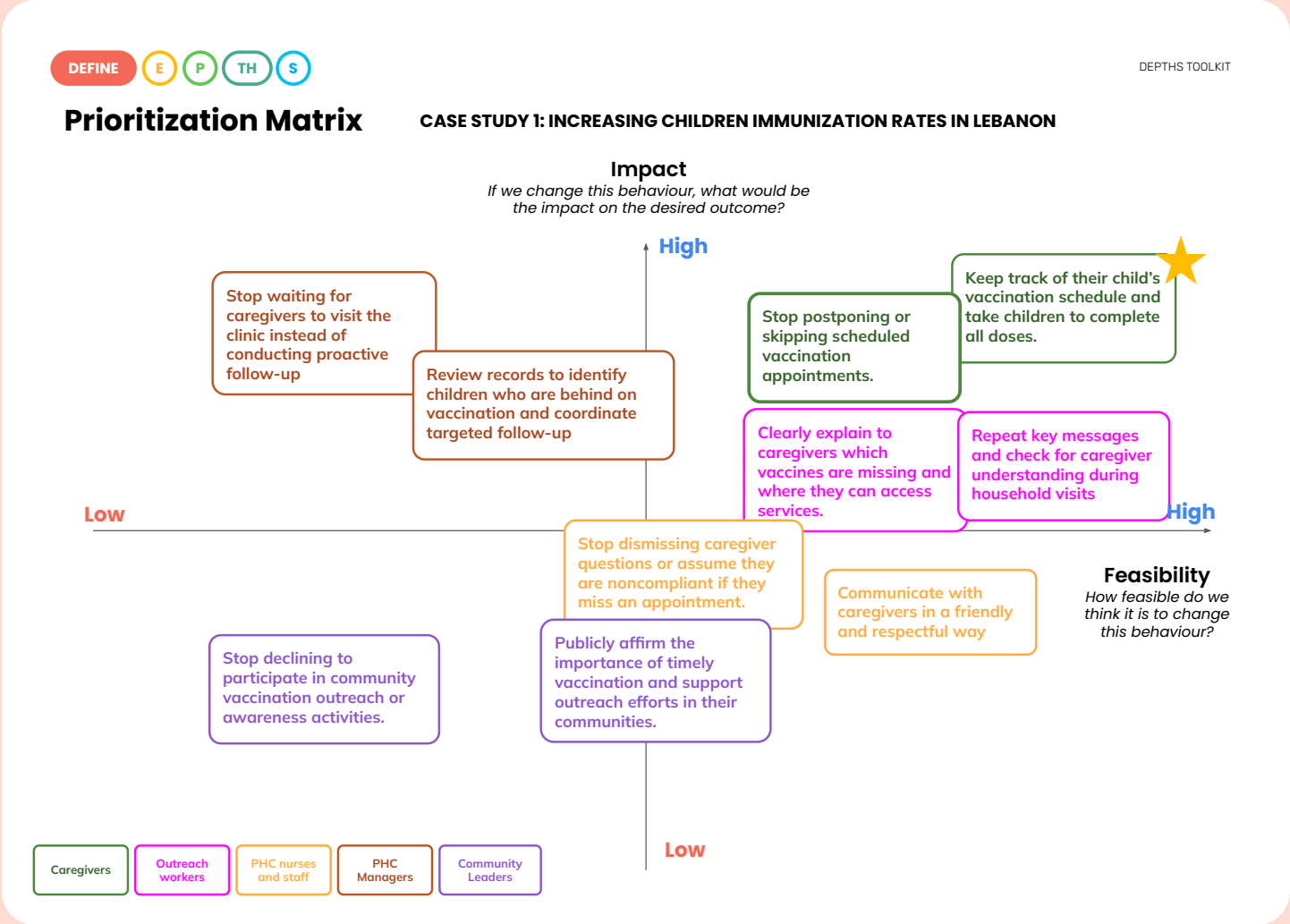
Increase timely completion of routine childhood vaccinations among children who are currently un- or under-vaccinated

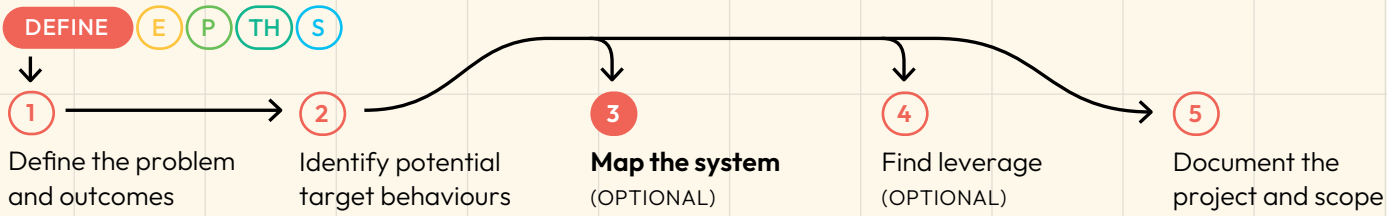
- **Caregivers**
  - **start** Proactively keep track of their child's vaccination schedule and take children to complete all doses.
  - **stop** Postpone or skip scheduled vaccination appointments.
- **Outreach Workers**
  - **start** Clearly explain to caregivers which vaccines are missing and where they can access services.
  - **start** Repeat key messages and check for caregiver understanding during household visits
- **PHC Nurses and Staff**
  - **start** Communicate with caregivers in a friendly and respectful way
  - **stop** Dismiss caregiver questions or assume they are noncompliant if they miss an appointment.
- **PHC Managers and District Supervisors**
  - **start** Review records to identify children who are behind on vaccination and coordinate targeted follow-up
  - **stop** Wait for caregivers to visit the clinic instead of conducting proactive follow-up
- **Community Leaders**
  - **start** Publicly affirm the importance of timely vaccination and support outreach efforts in their communities.
  - **stop** Decline to participate in community vaccination outreach or awareness activities.

With the behaviour map in hand, the project team used the Prioritization Matrix to weigh each behaviour by its potential impact on routine immunization and the feasibility of change within project constraints.

One key behaviour emerged as both impactful and feasible: caregivers keeping track of their child's vaccination schedule and taking children to complete all doses. Supporting this planning behaviour could help transform intention into action, particularly where services were available but underused.

The team treated the matrix as a guide not a rigid decision-making tool. They recognized that further research might uncover new factors or more nuanced behavioural barriers. The **Behaviour Tree** and **Prioritization Matrix** together offered a structured but flexible way to ground the project in real-world actions, creating a stronger basis for targeted interventions.





STEP 3:
Map the system (OPTIONAL)

In this step:

This step will use the System Map worksheet to contextualize the problem by creating a visual diagram of:

- The elements influencing the outcome(s) of interest
- The relationships between these elements (e.g., reinforcing, blocking, or enabling)

The goal is to move beyond surface symptoms and clarify how forces like infrastructure, social norms, psychological barriers and institutional rules interact, so interventions can shift the system as a whole.

Associated tools:

- System Map

This step is optional.

So far, the focus has been on defining the problem and outcome, and identifying priority behaviours for each audience. Sometimes, this narrow lens can overlook system-level forces that shape behaviour.

Step 3: Map the system draws on systems thinking — an approach that identifies different parts of the problem, their interactions, and how they influence each other over time.

This step is optional, but valuable if you want to uncover hidden barriers, feedback loops, or enablers that may not be visible at the behavioural level. If ready to move ahead, skip to Step 5: Build the project canvas and continue using the prioritized behaviour(s) from Step 2.

Why it matters:

Context is everything. Behaviours don't happen in a vacuum. They're shaped by systems: rules, services, (dis) incentives, social norms and expectations, and much more. Mapping the environment around a behaviour helps to:

- Capture knowledge from different experts and stakeholders who may each only see *part* of the system
- Build consensus and a shared understanding across teams and partners
- Reveal hidden dynamics such as feedback loops, bottlenecks, or unintended consequences
- Shift conversations from blame to systems thinking to open up new possibilities for change

Without a clear view of the system, it's easy to mistake symptoms for root causes, or to design well-intended solutions that solve one problem but create another. A system map helps answer: What's really driving this outcome? How do all the parts interact? Where are the hidden levers for change?

About feedback loops

Feedback loops explain how a change in one part of a system circles back to influence the original variable again.

There are two main types:

- A **positive feedback** or Reinforcing loop (+): They push the system further in the same direction. This can amplify change by creating a self-reinforcing cycle that accelerates the original shift. (e.g., more parents vaccinate → trust in the health system grows → even more parents vaccinate).
- A **negative feedback** loop is a process through which a system responds to change by counteracting or reversing that change, helping the system maintain stability or return to its original state. For example, if a community begins to overuse a health service, wait times increase, which may discourage others from using it.

In feedback loops, one factor both causes and is caused by another factor and when mapped out, it's easier to see how influence flows in both directions. Recognising these loops early helps to avoid unintended consequences and spillover effects, and reveals where small shifts can have system-wide impact.

(See resources at the end of this chapter for more on feedback loops.)

How to do it:

1. Map factors

Start by returning to the anchor for system mapping: Your Outcome Statement from Step 1. Then start identifying the building blocks of the system – the behaviours and their drivers. These include:

a. **Behaviours**: from Step 2 (Behaviour Tree)

b. **Drivers**: psychological (beliefs, emotions, routines, or cognitive biases), socio-cultural (norms, status, peer pressure, or cultural expectations), and environmental factors (access to infrastructure, availability of services, or physical surroundings) shaping those behaviours

c. **External influences**: existing programmes (community mobilization, outreach, other government programs etc.), policies (mandates, legislation, existing partnerships etc.), or broader context (e.g. economic instability, geographic barriers, humanitarian emergencies, conflict, misinformation)

Use the **Systems Map 1: Factors** worksheet:

- List key behaviours in the first column
- Add underlying drivers (beliefs, norms, access, rules, etc.) in the next columns
- Record existing efforts (programmes, policies, contextual factors) in the final column

Tips while mapping out factors:

- ➔ Look beyond the individual: Behaviour is shaped by context (e.g., norms, services, or infrastructure), not just personal choices.
- ➔ Don't aim for "the perfect driver": If a driver might be influencing the behaviour, include it. The drivers will be refined later in the process.
- ➔ Anchor in the Outcome Statement from Step 1 to stay focused.

DEFINE E P TH S

DEPTHS TOOLKIT

System Map 1: Factors

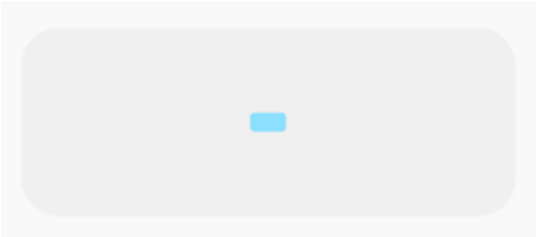
This step is best done on a large sheet of paper with sticky notes, or in a digital software like https://kumu.io/. List factors in the system: What makes it more likely, or less likely, that an outcome will happen?

Behaviours	Drivers	Existing programmes and efforts
What behaviours shape this outcome? *Refer back and pull from the Behaviour Tree. Name the stakeholders and the behaviours to explore further.*	***List the drivers that shape this outcome.*** *What influences people's decisions in this context - at the individual level, in their social environment, or in the context around them? What do people believe or feel about this behaviour? Are there habits or routines that shape what they do? What are the social norms, or pressures from family, peers, or the community? Are there practical enablers or barriers such as access, cost, time, rules, or infrastructure that affect whether the behaviour happens?*	***What is currently being done to address this problem?*** *Is there anything that drives the outcome that is outside our control?*

2. Build the system map

Next, bring the system to life visually using the **System Map 2: Map the System** worksheet (on a large piece of paper using sticky notes or with a digital tool like Miro, Kumu or Figma.[2])

a. Place the outcome statement (from Step 1) at the centre or top of the map, summarizing it in just a few words (e.g., 'increase timely completion of routine childhood vaccinations'). To make it stand out, place it in a shape, like a rectangle. Everything else on the map should connect back to this visualized outcome.



b. Add **key behaviours** (from Step 2 and also in System Map 1: Factors worksheet). Place them around the outcome and begin noting which stakeholders they relate to.

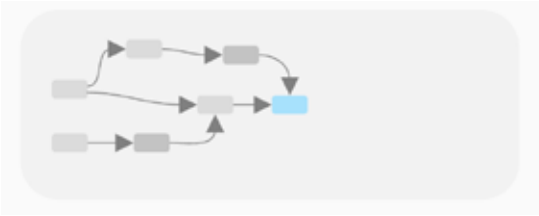
c. Link **drivers** (from the System Map 1: Factors worksheet) to the behaviours they influence with arrows.

d. Include **external factors and existing efforts**. Add in programmes, policies, services, and broader contextual factors that impact the drivers or behaviours from column 3 ('Existing programmes and efforts') of the System Map 1: Factors worksheet. Mark which ones are within scope and which are not.

Continues on next page...

2 These are collaborative digital tools that allow teams to create and edit visual diagrams in real time. They are ideal for mapping exercises.

e. Draw arrows to show **causal links** between elements. Show how one factor leads to or reinforces another. Ask: Does this factor affect that one? How? The arrows should point in the direction of influence, helping to visualize the causal chain.



f. **Add causal signs to the arrows** to show the type of effect each relationship represents and how the connections behave.

- (+) Positive: one factor increases the other (e.g., trust ↑ → vaccination ↑).
- (–) Negative: one factor reduces the other (e.g., wait times ↑ → satisfaction ↓)

g. Identify **downstream effects**. Ask: What happens if the outcome is achieved,or not? Ask:

- What longer-term or indirect effects might result from achieving this outcome?
- What risks or unintended consequences could emerge?

Add new boxes to the system map to represent each downstream effect. These effects could include impacts on health, learning, wellbeing, trust, productivity, or other areas.

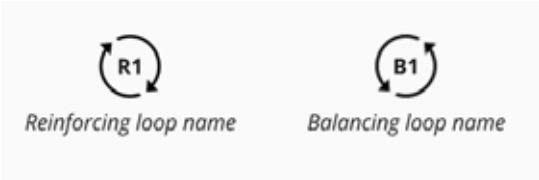
Next, connect each downstream effect to the outcome with an arrow that shows the direction of influence. Just like before, use causal signs and add:

- A green arrow or plus sign (+) if the outcome increases the downstream effect
- A red arrow or minus sign (–) if it decreases the effect

This step helps to visualize the broader implications of the team's outcome — and understand why it matters within the larger system.

Continues on next page...

h. Look for **feedback loops to highlight** when one factor influences another and that second factor, in turn, circles back to influence the first. This is where systems thinking becomes especially useful.



For example, a caregiver's distrust in the health system may lead them to delay or avoid vaccination. As fewer families participate in immunization, public health services become less visible or underutilized, which can reinforce perceptions that the system is unreliable, hence deepening the original distrust. In this example, labelling the feedback loop (e.g., 'trust erosion loop') can be helpful.

In the context of the case study on increasing timely childhood vaccination in Lebanon, consider this loop:

A caregiver delays taking their child for their routine vaccinations...

- ...increasing the chances of the child falling ill or missing scheduled doses...
- ...reinforcing the caregiver's belief that vaccination is complicated or not urgent...
- ...reducing their likelihood of responding to future reminders...
- ...therefore reinforcing the pattern of delayed or missed vaccinations.

Here, delayed vaccination both causes and is caused by negative perceptions and low follow-through, a classic negative reinforcing loop. Without breaking the cycle, the problem can intensify over time. Mapping these loops helps make it more clear where interventions can disrupt harmful cycles or strengthen positive ones.

i. **Consider differences across groups.** Ask: Is this experience the same for everyone? Are there historical disparities or social differences that might mean certain groups – like women, ethnic minorities, or rural populations — experience the system differently? Use notes, colour coding, or separate maps to highlight differences by gender, geography, ethnicity, disability, or other inequities Alternatively, a second version of the map that focuses solely on a particular subgroup can be developed. This can help ensure that analysis is inclusive and informed by principles of equity.

Continues on next page...

j. Finally, **share the systems map**. Treat the system map as a living document. Share it with colleagues, experts, and community stakeholders to check assumptions, fill gaps, and build shared ownership.

More information is available on system maps in the toolkit and training materials.

A few tips for the process of creating a system map:

- **Like any new skill, system mapping takes practice.** It may feel a bit messy or overwhelming at first, but it becomes easier (and the map becomes more insightful) with practice.
- **Focus on relationships, not just factors.** The power of a system map lies in how elements connect. Don't just list parts, map how they influence one another.
- **Don't aim for perfection.** The first map developed won't be the last. Think of it as a working draft that evolves as understanding grows.
- **Make it collaborative.** Mapping with others brings new perspectives, reveals blind spots, and builds shared ownership of the problem and its potential solutions.
- **Keep it visual.** Use colours, arrows and spacing to make relationships clear. If it's too dense to read at a glance, try simplifying or rearranging.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

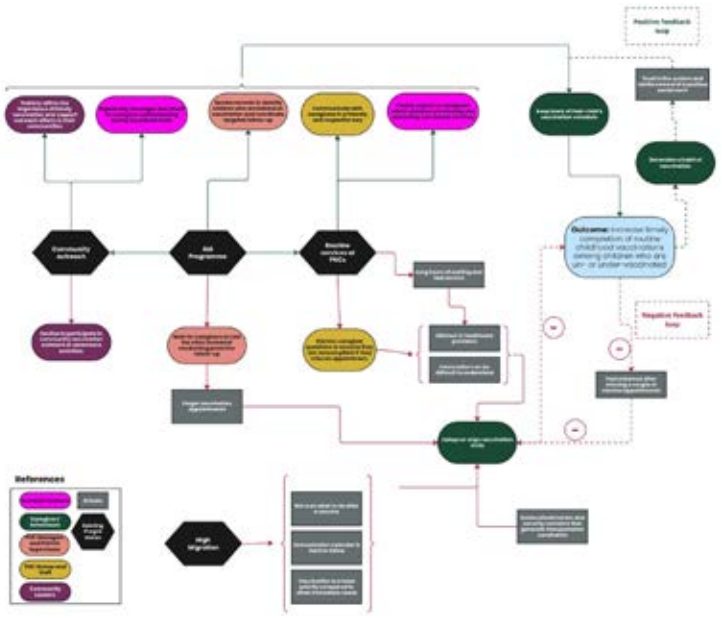
The System Map was not developed by the original project team. It is a recreated example based on real project data and context.

Before building the system map, the research team compiled a list of key behaviours, influencing factors (or drivers), and existing efforts. These components related to their outcome statement: increasing the timely completion of routine childhood vaccinations among children who are currently un- or under-vaccinated. They included:

- **Behaviours:** Most of the behaviours had already been identified and mapped in the **Behaviour Tree** exercise. These included actions such as caregivers responding to follow-up visits or Primary Healthcare Centre (PHC) staff reviewing vaccination records proactively.
- **Drivers:** Drawing on stakeholder conversations, prior existing data, and their experiences, the team listed potential drivers that might affect vaccination behaviours. For instance, they noted that some caregivers lacked trust in health workers, while others were constrained by time, mobility, or competing responsibilities. Additionally, social norms and misinformation shaped caregivers' beliefs about when, where and by whom children should be vaccinated.
- **Existing programmes and efforts:** As vaccination had long been a national health priority, the team considered ongoing initiatives, such as the Accelerated Immunization Activities (AIA) programme led by the Ministry of Public Health with UNICEF support. This programme included door-to-door outreach visits, referrals to PHCs, and caregiver education campaigns. The team also considered contextual factors like the refugee crisis, limited healthcare capacity, and pockets of low coverage identified in previous surveys.

Mapping behaviours

Starting with the desired outcome (i.e., the timely and complete vaccination of children) the project team added key behaviours from the **Behaviour Tree** exercise. They selected only the most relevant actions across the different audiences they had mapped, such as caregivers attending appointments, outreach workers providing referrals, and PHC staff updating records. The team maintained consistent colour coding to make the map easier to interpret and share with partners.



Mapping drivers

Next, the team layered in drivers, i.e., social, psychological, and environmental factors that influenced whether the behaviours occurred or not. These included:

- **Psychological drivers:** Caregivers underestimated the importance of timely vaccinations or assumed that minor delays were harmless.
- **Socio-cultural drivers:** In some communities, norms discouraged women from travelling alone, making it harder for mothers to take children to PHCs without accompaniment.
- **Environmental drivers:** Transportation barriers, long wait times at clinics, or unclear follow-up instructions created practical obstacles to vaccine completion.

These insights helped to reveal the range of factors shaping behaviour, beyond solely knowledge or access.

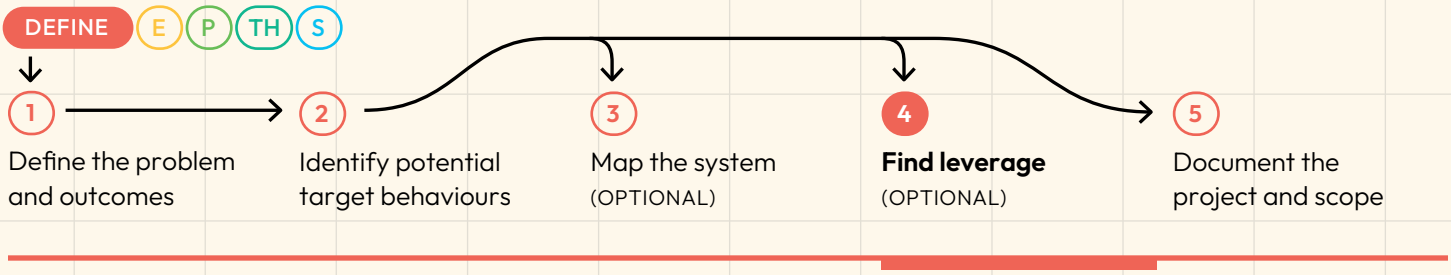
Mapping relationships and downstream effects

The system map also visualized how different behaviours and drivers were connected. For example, when nurses don't answer caregivers' questions in plain language, caregivers may feel confused. This confusion can lead to missed or delayed appointments. Over time, repeated missed visits may cause caregivers to feel ashamed and disengage from the health system altogether. Mapping these links helped the team to identify not only individual behaviours, but how the system either enabled or hindered change.

Feedback loops

Two key feedback loops emerged in the system. One was negative: caregivers who missed appointments felt confused or embarrassed, which lowered their likelihood of engaging with follow-up outreach. This effectively deepened mistrust and increased the chances of future missed vaccinations — causing a reinforcing cycle of disengagement.

The second loop was positive: when caregivers received friendly, timely follow-ups, and clear information, they were more likely to complete vaccination schedules. This strengthened their confidence in the health system, making them more likely to respond to future outreach and even encourage others — building a cycle of trust and engagement.



STEP 4:

Find leverage (OPTIONAL)

In this step:

In this step, you'll look at your system map to spot leverage points: places where change is both possible and powerful. For each, consider what behaviour could shift, what ripple effects it might create, and how realistic it is to act on. Leverage points are places in the system where small, strategic behaviour changes can unlock big improvements in outcomes. By analysing the system map, teams can identify leverage points and add them back into the map as strategic actions to guide the next phases of the project. This step helps ensure focus on a behaviour that is both impactful and realistic to implement.

Associated tools:

- Leverage Point Analysis

This step is optional.

Step 4 is also an optional step. However, if a system map was developed earlier (Step 3), it's strongly recommended to follow this step.

To 'Find leverage', the map is employed to identify leverage points. These are specific parts of the system where a strategic behaviour change could lead to a significant improvement in outcomes.

This step will help to:

- List promising behaviour changes that could shift key parts of the system
- Show ripple effects (i.e., how each behaviour change might influence other elements) across the system
- Assess feasibility by considering how realistic or actionable each leverage point is
- Restate the most promising leverage point and behaviour to change, which will guide the next phase: Explore and Diagnose

Why it matters:

Not every part of a system offers the same opportunity for change. Some elements are deeply entrenched or difficult to shift, while others, though seemingly small, can unlock significant changes in outcome. Those high-impact opportunities are known as leverage points. To identify leverage points, there are four key elements of analysis, which include:

- promising factors in the system
- potential behavioural changes related to those factors
- the ripple effects that such changes might generate across the system
- the feasibility of influencing those behaviours given the time, resources, and context

For instance, giving caregivers a personalised calendar with their child's photo and next vaccine date visibly marked may seem like a small change, but it can significantly improve timely immunisation by increasing salience, strengthening commitment, and providing a culturally resonant, low-cost prompt in the home.[3]

Identifying leverage points is how one moves from understanding a system to influencing it. With a clearer view of the key behaviours, drivers, and feedback loops, now begin to ask: Where could an intervention make the biggest difference?

Keep in mind: there is rarely one perfect solution. Instead, look for a few promising leverage points to explore and test further.

3 Abbott, P., Menzies, R., Davison, J. et al. Improving immunisation timeliness in Aboriginal children through personalised calendars. BMC Public Health 13, 598 (2013). https://doi.org/10.1186/1471-2458-13-598

How to do it:

1. Identify promising factors

Use the Leverage Point Analysis worksheet in the toolkit to help guide this reflection. Here's how to start:

a. Begin by revisiting the system map. Reflect on each factor – whether it's a behaviour, driver, or existing programme – and how it connects to others. Then, ask a few 'what if' questions to explore possible shifts:

- What would happen if this factor increased?
- What if it decreased?
- What if it disappeared completely, or connected to a different factor in a new way?
- Could a feedback loop be shifted to produce a better outcome?

For example, in the case of improving childhood immunization in Lebanon, factors to identify could include caregivers being more aware of their child's missed vaccinations or Primary Health Centre staff proactively following up with families who miss appointments. Each of these, if shifted, could trigger a positive ripple effect – such as improving trust, boosting follow-through, and ultimately increasing timely vaccine uptake. Focus on writing down factors that are considered both important and realistically influenced by one specific behaviour change.

b. Next, use the second column of the Leverage Point Analysis worksheet to write down potential behavioural changes that are related to the promising factors identified. These may be behaviours previously identified in Step 2 or new ones that emerged when reviewing the system as a whole.

This step takes the analysis deeper. It may confirm earlier behaviours or refine them. Changes might mean reducing or replacing a harmful behaviour, adding a new positive one, or scaling up an existing helpful one. Look for high-leverage points: areas where a small shift in behaviour can lead to a disproportionately large impact on the project's outcome. Don't worry about narrowing things down just yet – focus on capturing a range of possibilities.

What factor is promising? *On the 'Systems Map', ask what if each factor increases, decreases, disappears, or changes connection? Identify the promising factors.*	**What behaviour should change?** *List behavioural changes that could influence the promising factors. These may be new or previously identified. This is a deeper analysis.*
Promising Factors 1:	**Behaviour Change 1:**
Promising Factors 2:	**Behaviour Change 2:**
Promising Factors 3:	**Behaviour Change 3:**

2. Consider the ripple effects

Next, continue working on the **Leverage Point Analysis** worksheet by exploring how each potential behaviour change might ripple through the system.

a. For each behaviour change listed in the first column, ask: **What new outcomes might this trigger?** Could it cause downstream effects that amplify (or weaken) the original impact? How might it affect other parts of the system or other stakeholders over time? Write these reflections in the third column of the worksheet.

b. Rather than isolating each behaviour, **consider how a single shift could cascade across the system**, uncovering new opportunities or unintended consequences. The goal is to help compare and prioritize different leverage points based on their broader potential for impact.

For example, encouraging caregivers to proactively check their child's vaccination status and plan their visit in advance could have ripple effects such as improving the timeliness of vaccine uptake, reducing missed opportunities at clinics, and boosting overall trust in the health system. This could further ease pressure on outreach teams, who would spend less time tracking families for follow-up. The goal is to identify behaviour changes that don't just move the needle — they can multiply their impact across the system.

What factor is promising? *On the 'Systems Map', ask what if each factor increases, decreases, disappears, or changes connection? Identify the promising factors.*	**What behaviour should change?** *List behavioural changes that could influence the promising factors. These may be new or previously identified. This is a deeper analysis.*	**What are the ripple effects?** *What happens if this behaviour changes? How does that diminish or magnify over time and across the system?*
Promising Factors 1:	**Behaviour Change 1:**	**What are the ripple effects?**
Promising Factors 2:	**Behaviour Change 2:**	**What are the ripple effects?**
Promising Factors 3:	**Behaviour Change 3:**	**What are the ripple effects?**

3. Evaluate feasibility

The best leverage points balance impact and feasibility. First, ask:

- Is this behavioural change realistic within our context?
- Can the team or organization actually influence this behaviour?
- What kind of financial, technical, or human resources would be required to change this behaviour?
- How likely is it that this behaviour change will succeed?

Record the answers in the third column of the **Leverage Point Analysis** worksheet.

a. This assessment builds on the initial feasibility thinking introduced through the **Prioritization Matrix** in Step 2. At that stage, teams began identifying and comparing potential behaviours. Now, with deeper analysis from the system mapping and leverage point exercises, those early ideas can be revisited with greater insight. Some of the behaviours prioritized in Step 2 may still hold, while others may shift based on new insights.

b. Even if a change could be highly impactful, it may not be feasible in practice. For example, sending trained outreach teams door-to-door daily to remind caregivers of appointments might boost vaccine uptake significantly. However, doing so consistently at scale could require substantial staffing, funding and coordination. In short, the costs and operational burden might outweigh the benefits, and feasibility matters just as much as potential impact.

What factor is promising? *On the 'Systems Map', ask what if each factor increases, decreases, disappears, or changes connection? Identify the promising factors.*	**What behaviour should change?** *List behavioural changes that could influence the promising factors. These may be new or previously identified. This is a deeper analysis.*	**What are the ripple effects?** *What happens if this behaviour changes? How does that diminish or magnify over time and across the system?*	**Is it realistic?** *Can UNICEF drive this behaviour change? Is it cost-effective and feasible with time constraints and expertise?*
Promising Factors 1:	**Behaviour Change 1:**	**What are the ripple effects?**	**Is it realistic?**
Promising Factors 2:	**Behaviour Change 2:**	**What are the ripple effects?**	**Is it realistic?**
Promising Factors 3:	**Behaviour Change 3:**	**What are the ripple effects?**	**Is it realistic?**

4. Choose leverage points

After exploring several options, the next step is to select one leverage point, along with its associated behaviour change, that is most aligned with the team's goals, expertise, and operational constraints. This leverage point should sit in that 'sweet spot':, a behaviour that is both highly impactful and realistically achievable within our context.

Here's what to do:

a. The chosen leverage point may align with the target behaviour identified earlier using the **Prioritisation Matrix** in Step 2. If so, this reinforces the earlier decision. If not, that's perfectly acceptable — any differences can offer valuable insight. This step is designed to deepen understanding. A shift in focus, prompted by the system map or ripple analysis, may reveal a more promising entry point for intervention.

b. In order to choose the most promising leverage point, use the following set of criteria:

- **Feasibility:** Can we realistically influence this behaviour?
- **Impact:** If we succeed, how much would it move the needle on our outcome?
- **Equity:** Will this change promote a more equitable system or intervention?

c. Once a choice is made, return to the system map. Mark the leverage point clearly and highlight the behaviour it will target. Visually locating the intervention helps to clarify how it fits into the larger system and prepares the team for the next phase of work.

Before selecting leverage points, keep these tips in mind:

- **Avoid jumping to solutions.** This step is about understanding the system – not solving it yet. Focus on the behaviours (or lack thereof) that drive the problem, not on potential interventions.
- **Look carefully at behaviours that seem obvious.** We're all prone to availability bias – defaulting to what comes to mind easily. But 'familiar' does not necessarily mean 'successful'. It's important to explore a wider range of behaviour changes and evaluate them systematically.
- **Review existing evidence.** Before locking in a leverage point (i.e., target change in behaviour), check whether there is already research or experience showing what kinds of changes have worked in similar contexts. Build on what is known.
- **Consult subject matter experts and engage the community.** Talk to subject matter experts and community members: What would it mean to change this behaviour from your perspective?
- **Stay flexible.** The behaviour and leverage point chosen in this moment isn't final. It will continue to be validated and refined during the next phase, *Explore and Diagnose*. It's completely normal and encouraged for the focus to evolve as the team learns more.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The Leverage Point Analysis shown here is a recreated example based on real project data, used to illustrate what a completed tool could look like in practice.

Identifying promising factors and evaluating feasibility

The project team revisited their system map to examine the factors most strongly linked to their outcome: increasing timely completion of routine childhood vaccinations among un- or under-vaccinated children. For caregiver-focused behaviour, two key dynamics stood out. Caregivers often forgot appointments or deprioritised vaccines compared to urgent needs like food, work, or safety. These factors shaped real-world behaviour and were identified as promising to shift through targeted nudges or reminders.

For Primary Healthcare Centre (PHC) staff, the team identified a critical moment during clinic visits: caregivers were not always clearly told when to return or what vaccines remained. Although this information was part of the standard service, it was often skipped or inconsistently shared. This gap represented a missed opportunity to encourage proactive follow-through. Outreach workers in close contact with caregivers, often used unclear or overly technical messages. As a result, caregivers were left unsure what to do, where to go, or why completing the schedule mattered. These breakdowns in communication were identified as system weaknesses with high potential for improvement.

What factor is promising? *On the 'Systems Map', ask what if each factor increases, decreases, disappears, or changes connection? Identify the promising factors.*	**What behaviour should change?** *List behavioural changes that could influence the promising factors. These may be new or previously identified. This is a deeper analysis.*	**What are the ripple effects?** *What happens if this behaviour changes? How does that diminish or magnify over time and across the system?*	**Is it realistic?** *Can UNICEF drive this behaviour change? Is it cost-effective and feasible with time constraints and expertise?*
Promising Factors 1: Caregivers may forget vaccination appointments due to stress and competing priorities. Vaccination may be a low priority compared to urgent needs like food or work.	**Behaviour Change 1:** Caregivers build a consistent habit of attending vaccination appointments on time.	**What are the ripple effects?** Timely attendance improves vaccination completion rates, reinforces a culture of preventive health, and reduces the burden of follow-up on HCWs. Over time, this can increase trust in health services and normalize on-time vaccination in the community.	**Is it realistic?** Yes. This behaviour is already aligned with caregivers' goals but undermined by forgetfulness or unclear scheduling. It can be influenced through low-cost tools like reminders and planning supports, especially if paired with interpersonal reinforcement.
Promising Factors 2: The immunization calendar may be hard to follow. Caregivers may lack clarity on what to do after outreach	**Behaviour Change 2:** PHC staff routinely inform caregivers of the next vaccine date and encourage them to plan ahead.	**What are the ripple effects?** Clear next steps reduce drop-off between doses and increase caregivers' follow-through confidence and planning capacity.	**Is it realistic?** Yes. This fits within PHC staff's existing role. With supportive prompts, brief training, or standardized scripts, this behaviour can be adopted without major additional workload.
Promising Factors 3: Information may be misunderstood or not retained after visits. Repeated missed opportunities might discourage follow-through.	**Behaviour Change 3:** Outreach workers clearly explain the vaccination schedule and build urgency during home visits.	**What are the ripple effects?** Better messaging increases caregivers' understanding and perceived importance of timely vaccination, leading to more follow-through. Over time, this builds demand for vaccines and reduces the need for repeated visits and reminders.	**Is it realistic?** Yes, though it may require modest investment in job aids or training to ensure clarity, consistency, and empathy in outreach messaging. The role already exists and the change is an enhancement rather than an overhaul.

Evaluating leverage points and selecting promising behaviours

With these insights in place, the team identified three promising behaviour changes. One of the most impactful and feasible was supporting caregivers in developing a consistent habit of attending vaccination appointments on time. This change aligned with their existing intentions, but was undermined by forgetfulness or confusion. If reinforced with simple planning aids and trust-based reminders, it could raise timely vaccine completion and help normalize the practice across communities.

A second behaviour focused on PHC staff consistently informing caregivers about the next vaccine dose and encouraging them to plan ahead. This small, low-effort behaviour could yield significant benefits by reducing confusion and making follow-up easier, especially when paired with other outreach support.

Finally, the team highlighted a third behaviour: outreach workers clearly explaining the vaccine schedule during home visits and building a sense of urgency. This was viewed as a realistic improvement that could be supported through scripts, interpersonal training, or visual tools.

Mapping the leverage points in the system map

After identifying these key behaviours, the project team mapped them back into the broader system to understand their cascading effects. By improving caregiver planning and clarity, the team anticipated a reduction in missed appointments, increased trust in health services, and less strain on outreach and PHC staff. These changes could also create reinforcing feedback loops, such as caregivers gaining confidence and becoming role models in their communities.

To ensure relevance, the project team reviewed their findings with local implementers, outreach workers, and programme staff. This helped ground their leverage points in lived experience and align their strategy with operational realities. Ultimately, the process of identifying and embedding leverage points provided the team with a clearer, evidence-informed path for designing interventions that fit within the system — which had the potential to multiply impact across time.

Of the three behaviours, the team prioritized one: supporting caregivers to build a consistent habit of attending vaccination appointments on time." It was seen as highly impactful, but also realistically achievable with light-touch planning tools and clearer interpersonal messaging. It directly addressed one of the most immediate and widespread breakdowns in the vaccination journey, and aligned closely with the team's goal of improving timely vaccine completion among un- or under-vaccinated children.

STEP 5:

Document the project and scope

In this step:

This step describes how to complete the Project Canvas. This tool brings together all core project elements: outcome, target behaviour(s), stakeholders, scope, risks, early actions and next steps — into a single tool that anchors the project through DEPTHS.

Associated tools:

- Project Canvas

Why it matters:

The **Project Canvas** clarifies direction, aligns expectations, and surfaces any final assumptions or gaps. Without this step, teams often move forward with subtly different interpretations of what the project is aiming to achieve, leading to future misalignment.

This section also serves as the final checkpoint before moving into the second phase, Explore and Diagnose. Where possible, get input on the project canvas from the core team and key stakeholders which will strengthen buy-in, save time, and avoid any later confusion.

How to do it:

1. Build the project canvas

Find the Project Canvas worksheet in the toolkit.

a. Start by adding the project overview. List the project title, locations and description.

b. List the stakeholders. This includes those identified in Step 1.

- **Manage = Closely Involve.** These are key actors who are both influential and actively engaged. They should be closely involved in planning and decision-making.
- **Satisfy = Keep Engaged.** These stakeholders are influential, but they may not be actively engaged. Keep them informed and satisfied and find ways to maintain their support.
- **Inform = Consult and inform.** These actors care about the issue but have limited influence. Keep them informed and consult them for insights, especially those with lived experience.
- **Monitor = Track for later.** These groups may not be central to the current work, but they could become more relevant later. Keep an eye on their level of interest and influence over time.

c. Write down what's in and out of scope for the project. Include the DEPTHS phases involved, as well as the specific deliverables, such as research reports or prototypes. Each DEPTHS stage matters. Skipping steps, like moving ahead without diagnosing root causes or testing interventions, can lead to weak or ineffective outcomes. Occasionally, some phases can be adapted or condensed and shorten the process. Condense and adapt:

- The Define phase if the SMART outcome and audiences are already agreed on by partners in the last 6 months and grounded in recent programme data, skipping the Define phase could be possible.
- The Explore and Diagnose phase if there is already existing strong, recent, local evidence on the same behaviour and population, with clear behavioural drivers have been identified.
- The Prototype Designs phase if the scope is adapting a proven pattern that has already been used in very similar contexts.
- The Test Hypotheses phase if there are very low-risks of change or that only small operational tweaks with limited exposure are needed.
- The Scale phase when the prior phases have produced evidence in this context and implementation capacity exists.

d. Add risks, which are the anticipated challenges that may affect the success of the project. Risks may be:

- **Logistical:** such as access to communities during rainy season

- **Political:** such as government transitions or policy shifts
- **Financial:** such as limited budgets or funding delays
- **Operational:** such as staffing gaps or data security concerns
- **Behavioural:** such as resistance from key stakeholders or social norms

The idea isn't to predict everything, but to think about key issues and, where possible, how to mitigate them.

e. List outcome(s): the real-world change the project is aiming to achieve. Pay special attention to change that impacts people's lives.

f. Add the project's audiences. This is the community or group(s) whose behaviours are the focus of the project's work. In some cases, this may be the group you actively engage or co-create with, such as caregivers or adolescents.

g. Write down the behaviours to explore. This will include priority behaviours identified in earlier steps that are most likely to influence the outcome.

DEFINE E P TH S

DEPTHS TOOLKIT

Project Canvas

Fill out the canvas at the start of the project by pulling information from the previous activities. As we learn more, update as needed.

Overview

PROJECT TITLE:

LOCATIONS:

DESCRIPTION:

Stakeholders

MANAGE:

SATISFY:

INFORM:

MONITOR:

Outcome

Real world change: *What does this project accomplish?*

Audiences

Community of focus: *Who do we engage, collaborate with, and deliver this change to?*

Scope

STAGES

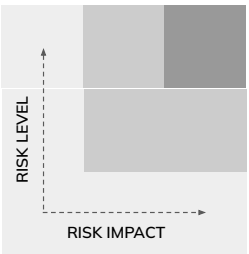
DEFINE | EXPLORE | PROTOTYPE | TEST HYPOTHESES | SCALE

SIMILAR / PREVIOUS PROJECTS

DELIVERABLES

Risks

What challenges do we anticipate? How can this be avoided or reduced?



Behaviours to explore

List the primary and secondary audiences and the behaviours to explore for each. *These are the priority behaviours identified in earlier steps, those most likely to influence the outcome we're aiming to shift. Behaviours should be observable, specific, and have a clear link to the outcome we're aiming for.*

Keep these tips in mind while building the Project Canvas:

- **Consider how different groups will be affected.** In behavioural science, the behaviour of interest will sit at the individual level among the populations we aim to serve. However, if the change places too much burden on already marginalized groups, pause and reconsider. Go back to the Behaviour Tree or System Map to see if the responsibility could be shared more equitably through systems, services, or norms. If so, pause. Revisit the Behaviour Tree or System Map and reconsider selected leverage points, key actors, and priority behaviours.
- **Consider both individual and communal dynamics.** Identify the behaviours of multiple actors, institutions, service providers, community groups, or families. This helps to uncover bottlenecks and find possibilities for system-wide change.

2. Expand the team

After an initial assessment of the problem and a clearer understanding of the context, it might be necessary to recruit new members for the project team. Identify any skill gaps and roles needed, and look for additional team members that could cover those gaps, ensuring the team is well positioned to carry the project through the next phases.

Throughout DEPTHS, three core technical areas are woven together: behavioural science, human-centred design, and systems thinking.The team doesn't need to have deep expertise in each of these three areas, but having some experience or familiarity with each will improve the ability to move through the process effectively. In some cases, different phases of DEPTHS may be led by different teams: a research team may handle the 'DEP' side, while an implementation team handles the 'THS' side. As such, it's important to clarify which steps of the DEPTHS process the current project will cover, and if different teams are involved, ensure handovers are well managed.

The backgrounds and experiences of the team members are equally important and valuable. Including local stakeholders and members of the target community, either directly on the team or through an advisory board, helps to ensure that the work is grounded in a real-world context, and truly reflects the voices of those most affected.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The Project Canvas shown here is a recreated example based on real project data, used to illustrate what a completed tool could look like in practice.

At this stage of the project, the team had developed an outcome statement, a behaviour tree, a prioritization matrix, a system map, and a leverage point analysis. Based on these steps, they decided to prioritize changing caregivers' behaviour: to support caregivers in building a consistent habit of completing their children's vaccinations on time. The aim was to reduce delays and drop-offs in routine childhood immunizations among refugee and host community households identified as un- or under-vaccinated.

Overview

The canvas included the project title, geographic scope (three districts in Lebanon), and population focus (un- or under-vaccinated children aged 0–16 years). Key institutional stakeholders included the Lebanese Ministry of Public Health (MoPH), UNICEF, and affiliated outreach teams conducting household visits. Anticipated risks ranged from caregiver mistrust in the health system to logistical challenges in sustaining household follow-up.

Outcome

The outcome statement centred on increasing timely completion of routine childhood vaccinations among un- or under-vaccinated children.

Community of focus

The primary agents of change were caregivers — in particular, mothers and fathers in refugee and low-income host communities. These were the individuals most directly involved in vaccination decisions. System mapping confirmed they were the audience with behaviour that was the most impactful and the most feasible to influence. Outreach workers and PHC staff were also seen as key players in the system, but their behaviours were deprioritised at this stage, in favour of focusing on caregiver action.

Overview

PROJECT TITLE: Increasing Timely Childhood Vaccination in Lebanon

LOCATIONS: 3 districts in Lebanon with low immunization coverage.

DESCRIPTION: Improve timely completion of childhood vaccinations through a behavioural science approach.

Stakeholders

MANAGE: The technical project team within the MoPH

SATISFY: High level public officials and donors

INFORM: Caregivers of young children, including both Lebanese and Syrian refugee families.

MONITOR: Local NGOs

Outcome

Increase timely completion of routine childhood vaccinations among children who are currently un- or under-vaccinated in targeted areas and targeted by Lebanon's Accelerated Immunization Activities (AIA) programme.

Audiences

Caregivers of un- or under-vaccinated children: Primarily mothers, but also fathers, grandparents, or other household members responsible for child health.
Un- or under-vaccinated children (ages 0–16): Children identified as behind on Lebanon's routine immunization schedule.

Scope

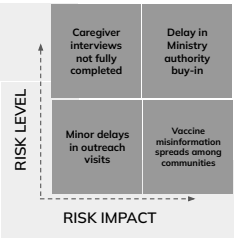
STAGES

DEFINE | EXPLORE | PROTOTYPE | TEST HYPOTHESES | SCALE

SIMILAR / PREVIOUS PROJECTS
Promoting follow-up attendance among mental health patients at a primary healthcare center in Lebanon (Kontar et al., 2022)

DELIVERABLES: Final report (presentation and academic paper-style document)

Risks
What challenges do we anticipate? How can this be avoided or reduced?



Behaviours to explore

Caregivers build a consistent habit of attending vaccination appointments on time.

PHC staff routinely inform caregivers of the next vaccine date and encourage them to plan ahead.

Outreach workers clearly explain the vaccination schedule and build urgency during home visits.

Actors and actions

Through the Behaviour Tree and Leverage Point Analysis, the team identified three main behaviours contributing to vaccination delays, with one chosen as the top priority. The central behaviour was as follows: caregivers forming a clear and timely vaccination plan, following through on appointments, and avoiding unnecessary delays.

Risks

As the project team prepared to move into the next phase, they identified several key risks.

- The most critical – high risk and high impact – was a potential delay in securing buy-in from Ministry authorities, which could stall implementation or limit institutional support.
- A second concern, with high risk but lower impact, was that some caregiver interviews might remain incomplete, affecting data comprehensiveness but not derailing the project.
- Among lower-probability risks, vaccine misinformation spreading within communities was flagged as low risk but high impact, given its potential to undermine trust and uptake.
- Lastly, minor delays in outreach visits were considered low risk and low impact, and likely manageable through scheduling adjustments.

Final checklist for *Define*

- [ ] Problem Definition and Outcome Statement
- [ ] Stakeholder and Audience Map
- [ ] Behaviour Tree
- [ ] Prioritization Matrix
- [ ] Project Canvas

Optional:

- [ ] System Map
- [ ] Leverage Points

Learn more

This field guide is designed to equip teams with practical tools, frameworks and methodologies to apply behavioural science to a range of real-world challenges. As behavioural science draws from multiple disciplines — including human-centred design, experimental economics, and systems thinking — we've curated a selection of approaches that reflect this diversity. The following section offers additional resources to explore specific topics introduced in the guide, along with the option to continue a self-paced learning journey.

"I want to learn more about human-centred design and how to co-create with my community."

There are many valuable resources available to help meaningfully engage with communities and tap into local expertise. If looking to involve community members in identifying and prioritizing solutions, the field of human-centred design (HCD) offers helpful starting points. Within UNICEF, you can explore the SBC Guidance for HCD and the HCD Field Guide, both tailored to support practitioners working in diverse settings.

Curious to learn more from outside sources? Consider IDEO's Field Guide to Human-Centred Design. As one of the pioneers in formalizing HCD practices, IDEO offers accessible, practical guidance drawn from years of experience co-creating solutions around the world.

"I want to explore additional case studies."

Beginning this journey in-house is always a great option. UNICEF has an extensive variety of publications and research, as well as tools designed to support children across its five core programmatic areas. If interested in exploring how behavioural science is applied beyond UNICEF, there are many other organizations doing impactful work in this space.

Organizations that regularly share their insights and methods include Common Thread, Busara, FirstHand, the Behavioural Insights Team, Ideas42, J-PAL, the World Bank's eMBeD, and the Inter-American Development Bank, among others. Reviewing their work can offer fresh perspectives and practical examples to enrich an applied behavioural science approach.

"I want to learn more about systems thinking and mapping."

One of the best resources to learn more is the open access book Systems Mapping, which provides an excellent non-technical summary of seven different approaches.

In particular, the approach here builds on the standard technique of causal loop diagrams, with a focus on behavioural factors. A good guide to causal loop diagrams can be found here.

Another excellent resource comes from the System Mapping Academy's online materials, including a free Toolkit. For approaches that are particularly useful for applied behavioural science, see UCL's behavioural system mapping and Busara's behavioural systems analysis. This online summary provides links to many other resources.

"I want to learn more about how to find good leverage points that might have significant ripple effects."

For further insights into ripple effects and leverage points, the scientist, writer, and educator Donella Meadows gave a celebrated and highly-cited talk on systems thinking. Meadows' talk concerns the multiple levels to employ while considering systems change, including changes in the underlying function or purpose of a system.

Resources

1. Behavioural Insights Team (BIT), *Target, Explore, Solution, Trial, Scale: An introduction to running simple behavioural insights projects*, 2022, <www.bi.team/wp-content/uploads/2022/11/BIT-Handbook-How-to-run-simple-BI-projects.pdf>.
2. Henrich, Joseph, 'WEIRD', *Open Encyclopedia of Cognitive Science*, edited by Michael C. Frank and Asifa Majid, MIT Press, 24 July 2024, <https://doi.org/10.21428/e2759450.8e9a83b0>.
3. IDEO.org, *The Field Guide to Human-Centered Design*, <https://www.designkit.org/resources/1.html>, accessed 12 June 2025.
4. IDinsight, 'Evidence Reviews', <https://guide.idinsight.org/evidence-reviews/>, accessed 12 June 2025.
5. Landry, Lauren, 'What is Human-Centered Design?', Harvard Business School Online, 25 November 2020, <https://online.hbs.edu/blog/post/what-is-human-centered-design>.
6. Meadows, Donella, 'Leverage Points: Places to Intervene in a System', Donella Meadows Project, https://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/>, accessed 12 June 2025.
7. Meadows, Donella, *Thinking in Systems: A Primer*, edited by Diana Wright, Chelsea Green Publishing, December 2008, <www.chelseagreen.com/product/thinking-in-systems/>.
8. Miro. 'System Mapping Toolkit', https://miro.com/miroverse/system-mapping-toolkit/, accessed 12 June 2025.
9. Smaldino, Paul E., *Modeling Social Behavior*, Princeton University Press, October 2023, <https://press.princeton.edu/books/paperback/9780691224145/modeling-social-behavior>.
10. UNDP, *System Change: A Guidebook for Adopting Portfolio Approaches*, 28 March 2022, <www.undp.org/publications/system-change-guidebook-adopting-portfolio-approaches>.
11. UNICEF Social and Behaviour Change (SBC) Guidance, 'Using Human-Centred Design', <www.sbcguidance.org/create/using-human-centred-design>, accessed 12 June 2025.
12. UNICEF Social and Behaviour Change (SBC) Guidance, 'Social and Community Listening', <www.sbcguidance.org/do/social-and-community-listening>, accessed 12 June 2025.
13. University of California Office of the President, 'How to Write SMART Goals: A How to Guide', <www.ucop.edu/local-human-resources/_files/performance-appraisal/How+to+write+SMART+Goals+v2.pdf>, accessed 12 June 2025.

Explore & Diagnose

Welcome to *Explore & Diagnose*!

Gathering the necessary evidence to inform intervention design is the critical next step to applying behavioural science.

This phase focuses on uncovering behavioural drivers and barriers through the use of research and data. It involves synthesising existing social and behavioural evidence, identifying gaps where additional research is needed, and building a clear picture of the problem — all in order to guide where and how to intervene.

Why *Explore & Diagnose*?

Even with a clear problem and outcome in mind, it's easy to make assumptions about what is driving people's behaviours or what might change them.

However, behaviours are rarely shaped by a single factor. They are influenced by a wide range of forces, from beliefs and habits to social norms and structural conditions. The *Explore & Diagnose* phase allows researchers to understand the context, surface hidden barriers, and use this knowledge to design relevant solutions.

How can we *Explore & Diagnose* the challenge?

In this second phase of the DEPTHS methodology, there are five steps to guide teams through the behavioural research process.

Each step includes specific tools, guidance on why these tools matter, and how to use them. A **case study on increasing childhood vaccination uptake in Lebanon** will exemplify how the tools can be applied in practice.

Before starting *Explore & Diagnose*, make sure to complete all the steps of the Define phase.

Summary of the *Explore & Diagnose* phase

Before you start: review the major *Define* phase outputs: Problem Definition and Outcome Statement, Stakeholder Map and Target Audience, Behaviour Tree, Prioritization Matrix, System Map (optional), Leverage Point Analysis (optional), Project Canvas

1 Conduct initial desk research	2 Develop primary research questions	3 Plan the primary research	4 Perform the primary research	5 Analyse and synthesize the findings
Review existing data and literature to build a foundational understanding of the context. This helps to surface known insights, identify gaps, and determine whether additional research is needed.	Define clear and behaviourally informed research questions aligned with gaps identified during the desk review. Questions should focus on uncovering why behaviours do or do not occur, using behavioural frameworks like COM-B.	Outline the research approach, including who to engage, what methods to use, and how ethical safeguards will be applied. This step ensures that research is practical, targeted, and ethically grounded.	Organize and implement fieldwork, including participant recruitment, team training, tool testing, and logistics. Carry out data collection while emphasizing quality, inclusion, and respectful engagement.	Use the COM-B framework to organize insights, develop behavioural profiles, map micro-behaviours, and prioritize key barriers and enablers that will shape the next stage of design.
Initial Desk Research	Primary Research Objectives and Questions	Primary Research Plan [External] Research Protocol Template		Behavioural Profile Behavioural Map and Diagnosis Feasibility-Impact Matrix

Common Pitfalls

Common traps can derail behaviourally informed work, so throughout the *Explore and Diagnose* phase, keep these possible issues in mind:

- **Skipping or rushing the Define phase review.** Jumping straight into research without revisiting outputs like the Problem Statement, System Map, or Project Canvas often leads to duplication, misalignment, or unclear research objectives. For example, a team exploring low immunization uptake might conduct interviews about vaccine beliefs, without noting that the original Prioritization Matrix has already identified caregivers proactively keeping track of their child's vaccination schedule as the central, prioritized target behaviour.
- **Desk research should go beyond a quick literature scan.** This process requires a structured approach, including reviewing diverse sources (e.g., academic, programmatic, behavioural, and policy data), asking targeted questions about the behaviour of interest, and synthesising insights and what is currently known through a behavioural lens. Reviews that don't follow a structured approach may skip critical sources, or fail to acknowledge existing findings, leading to poorly scoped primary research. For example, if a team exploring low immunization uptake failed to analyse caregiver perception surveys, this might result in overlooking existing concerns about provider trust, which could skew design priorities.
- **Writing broad or vague research questions.** If research questions aren't specific to the behaviour of interest, they can result in generic findings that are difficult to act upon. Each question should be linked to a decision. For example, asking "What are barriers to vaccination?" is too broad; "What makes caregivers delay the second dose after the child's first clinic visit?" is clearer and more actionable.
- **Not aligning on whose voices matter.** Failure to identify and prioritize actors such as caregivers, influencers, or service providers can result in missing critical perspectives. For example, a project might focus only on caregivers but miss the influence of grandmothers or community health workers, who shape vaccination decisions. Instead, map and prioritize actors early.
- **Analysing data without unpacking the behavioural drivers behind it.** Too often, findings are catalogued at face value without unpacking the underlying drivers, barriers, or heuristics that shape people's decisions. For example, noting that "caregivers fear side effects" is only useful if we push further, understanding where the fear comes from, and identifying both who or what could shift it. Without this, insights remain surface-level and difficult to act upon.
- **Treating all barriers as equally important.** Without prioritization, teams may spread their efforts too thin. This stage of the *Explore & Diagnose* phase concerns finalizing the research process by prioritizing the behavioural barriers that matter most: those with the greatest influence on the target behaviour and the highest potential to be addressed through the design process. For example, spending equal time on minor myths and major structural barriers — like clinic distance — could dilute resources by splitting efforts across challenges that don't have equal weight. Using a Feasibility–Impact Matrix can help teams to focus efforts accordingly.
- **Jumping to solutions too early.** The goal of this phase is to gain a deep understanding of behaviours, not to generate or test solutions. Premature ideation can limit curiosity, bias evidence interpretation, and result in interventions that are misaligned with real user needs. For example, suggesting reminder apps before assessing digital access might overlook that most caregivers do not use smartphones — leading to a solution that is irrelevant or inequitable.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

Childhood vaccination is one of the most cost-effective public health interventions, yet many children, especially in low- and middle-income countries, remain un- or under-vaccinated.

In Lebanon, home to the world's highest per capita refugee population, vaccine uptake has been undermined by poverty, displacement, and strained health systems. Although national coverage once approached 90%, outbreaks of measles and mumps revealed growing pockets of under-vaccination. In response, the Ministry of Public Health and UNICEF launched the Accelerated Immunization Activities (AIA) programme to expand access through health centers and community outreach — but uptake remained low.

A multidisciplinary team from UNICEF, Nudge Lebanon, and MoPH set out to investigate the issue through a behavioural science lens. Rather than focusing solely on structural barriers, they examined specific behaviours and their drivers, with a focus on caregiver decision-making. During the Define phase, the team articulated a clear outcome: Improving the timely completion of routine childhood vaccinations among un- or under-vaccinated children. They mapped the ecosystem of actors, including caregivers, outreach workers, and PHC staff. They used behavioural mapping and a prioritization matrix to identify "start" and "stop" behaviours across these audiences, ultimately prioritizing caregiver follow-through on appointments as both impactful and feasible to shift.

Optional steps included a system map and leverage point analysis, which revealed how forgetfulness, unclear instructions, and low perceived urgency contributed to missed vaccinations. By Step 5 of the *Define* chapter, the team had documented a clear project scope, surfaced potential risks, and was ready to move into deeper exploration of behavioural drivers. This case study will continue to serve as an illustrative example throughout the Explore and Diagnose phase, showing how behavioural insights can be applied to guide each step of the process.

The team began by identifying districts with the lowest coverage rates and the most vulnerable populations, focusing on areas with a high concentration of refugee and low-income households. Using behavioural science frameworks — such as the **COM-B**[1] framework to explore capability, opportunity, and motivation — they conducted extensive fieldwork, including interviews with outreach workers, program staff, and healthcare providers, along with direct observation through household visits. They mapped the journey caregivers were expected to take and looked closely at the specific decision points where behaviour broke down, without making assumptions about lack of awareness or access.

This structured approach revealed clear behavioural drop-off points, i.e., moments where intention failed to translate into action. Three key barriers emerged:

- **Cognitive overload**: caregivers feeling overwhelmed by too much information or too many competing tasks
- **Emotional stress**: such as anxiety about clinic visits or fear of judgment
- **Social perceptions:** e.g., concerns about how others in the community might view their actions

These barriers did not reflect an opposition to vaccination but rather the complex realities that caregivers navigated, where attending a clinic often competed with urgent responsibilities like securing food, childcare, or income. By grounding the inquiry in observed behaviour and real-world frictions, the team was able to diagnose where and why caregivers' intentions were breaking down, often in ways that were unconscious or invisible even to the caregivers themselves. These were not always moments of deliberate decisions; rather, they reflected subtle, automatic behavioural responses shaped by stress, habit,

1 Michie, S., van Stralen, M.M. and West, R., 2011. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science, 6(1), p.42. https://doi.org/10.1186/1748-5908-6-42

and environmental context. Caregivers might not have felt like they were facing barriers, but the behavioural patterns revealed key friction points where intention to follow-through on vaccination quietly unravelled.

***Note:** While this is a real project that closely followed a very similar process to DEPTHS, there were a few tools from the toolkit that the project team did not apply during implementation. In those cases, we've gone back and retrospectively applied the tools using real project data to illustrate how they might have looked if they had been used at the time.*

STEP 1:
Conduct initial desk research

In this step:

The purpose of this step is to conduct an initial desk research (sometimes referred to as a “literature review”).

This is a structured process of examining and synthesising existing data and evidence.

Associated Tools:

- Initial Desk Research

Why it matters:

Before conducting primary research, it is essential to understand the existing landscape of evidence.

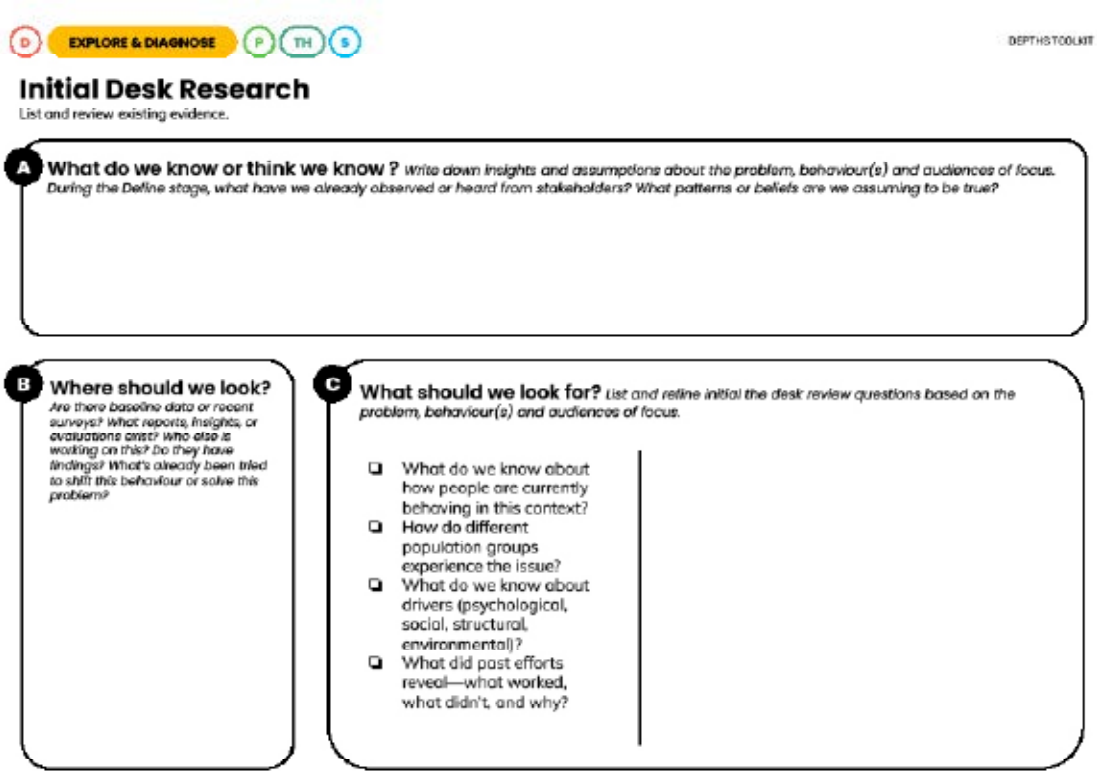
A rigorous desk research helps to build a shared understanding of the behavioural challenge, identify well-documented behavioural patterns and influencing factors, and pinpoint clearly where further information or research is still needed. If these gaps are significant, the desk research can shape a focused plan for primary research, ensuring any new data collection is purposeful and targeted to the areas that matter most.

How to do it:

1. Review the outputs from the *Define* phase and formulate initial questions.

Use the Initial Desk Research worksheet to prepare the desk research process.

a. Before initiating the desk research, revisit the key outputs generated during the *Define* phase, including the Problem Definition and Outcome Statement, Behaviour Tree, Prioritization Matrix, System Map, Leverage Point Analysis (if available) and the Project Canvas. This step helps the team to consolidate what is already known, clarify working assumptions, and sharpen the focus of the desk research.



D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Initial Desk Research

List and review existing evidence.

A What do we know or think we know ? *Write down insights and assumptions about the problem, behaviour(s) and audiences of focus. During the Define stage, what have we already observed or heard from stakeholders? What patterns or beliefs are we assuming to be true?*

B Where should we look? *Are there baseline data or recent surveys? What reports, insights, or evaluations exist? Who else is working on this? Do they have findings? What's already been tried to shift this behaviour or solve this problem?*

C What should we look for? *List and refine initial the desk review questions based on the problem, behaviour(s) and audiences of focus.*

- ❑ What do we know about how people are currently behaving in this context?
- ❑ How do different population groups experience the issue?
- ❑ What do we know about drivers (psychological, social, structural, environmental)?
- ❑ What did past efforts reveal—what worked, what didn't, and why?

b. Complete the "What do we know or think we know?" section of the **Initial Desk Research** worksheet:

- What insights have already been established through observation or stakeholder input so far?
- What assumptions emerged during *Define* that need to be validated through further research?

c. Complete the "Where should we look?" section. Identify relevant sources for the desk research, such as:

- Programme evaluations and implementation reports
- Administrative data
- National demographic and health surveys (e.g., DHS, MICS)
- Government policy documents and strategic plans
- Academic literature and peer-reviewed studies
- Grey literature and unpublished reports
- Internal data, learning briefs, or field-level observations from partners

Draw on both qualitative and quantitative data — from local and regional sources — to build a comprehensive understanding of the context. Prioritize recent, credible, and context-specific evidence that speaks to both the situation and the behaviour. This includes essential data on service coverage, access, and demographics, as well as sources that help to explain what drives people's decisions.

Review data from national statistical offices, UN agency databases, Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), national health information systems (HMIS), and sector-specific reports from ministries or NGOs. For insights on behavioural drivers and barriers, review existing KAP surveys, social listening reports, and rapid qualitative assessments conducted in similar contexts.

Combining situational and behavioural evidence provides a stronger foundation for analysis — helping teams to move from understanding what is happening to unpacking why it is happening.

d. Finally, use the "What should we look for?" box to tailor and refine initial desk research questions. Prompts to consider when formulating initial desk research questions include:

- What do we know about how people are currently behaving in this context?
- How do different population groups experience the issue?
- What do we know about drivers (psychological, socio-cultural, structural, environmental)?
- What did past efforts reveal? What worked, what didn't, and why?

2. Conduct the initial desk research

a. Review and analyse available literature, drawing from both situational and behavioural evidence. When doing so, pay particular attention to the *behavioural* insights. What do the sources say about what people are doing and why? What barriers or enablers are noted? These may be explicitly mentioned, or implied instead through narrative or outcome patterns.

b. Use different search engines to look for past evidence. In addition to general academic search engines like Google Scholar, consider more specialised platforms to refine desk research like the 3ie Development Evidence Portal, UNICEF's publications or the Behavioural Evidence Hub. AI-powered tools, such as the World Bank's AVA chatbot, and tools for literature reviews like Elicit, can support an efficient discovery of evidence.

Whatever the tool, take care to check primary sources identified through AI tools to ensure accurate context and minimize bias.

c. It is important to critically assess the literature and evidence collected during a desk research. Different journals and platforms apply varying standards for publication, and even peer-reviewed articles may not always be reliable.

Not all evidence carries the same weight. The following table offers guidance for weighing different kinds of evidence.

Where possible, the desk research should prioritize evidence from systematic reviews, meta-analyses, and other high-quality studies, if these are available.

TABLE 2. TRUSTWORTHINESS ACCORDING TO TYPE OF EVIDENCE

EVIDENCE TYPE	TRUSTWORTHINESS
Systematic reviews and meta-analyses of high-quality studies	**Very strong**
Multiple RCTs or well-designed quasi-experimental studies	**Strong**
Single RCTs, strong observational studies, mixed-methods triangulated evidence	**Moderate**
Case studies, cross-sectional surveys, grey literature with limited rigour	Weak
Anecdotes, expert opinion, advocacy briefs with no data	Very weak

The UK Department for International Development (DFID) provides guidance on evaluating the strength of evidence, as illustrated below in the table below:

TABLE 3. PRINCIPLES OF HIGH QUALITY RESEARCH STUDIES

CRITERIA	DESCRIPTION	HIGH QUALITY EVIDENCE EXAMPLE	LOW QUALITY EVIDENCE EXAMPLE
Conceptual framing	The study explains its main question and links it to what is already known.	A report on SMS reminders for vaccines states its research question clearly, cites past evidence on reminders, and sets out a testable hypothesis.	A paper simply says "we wanted to improve vaccination" without explaining why, or linking to earlier studies.
Transparency	The study is open about how it was done (where, when, with whom, and who paid for it.)	A malaria study explains it was conducted in 3 provinces in 2022, with 600 households randomly chosen, and names UNICEF as funder.	A study reports "data were collected recently in Africa" with no details on location, methods, or funding.
Appropriateness	The way the study was done fits the question being asked.	To measure whether a new drug prevents malaria, researchers run a randomised controlled trial.	To test if a drug prevents malaria, researchers just ask 10 people if they feel healthier after taking it.
Cultural sensitivity	The study considers local customs, beliefs, and context.	A vaccination survey in Nigeria notes that fathers usually decide on children's vaccines, so it includes both mothers and fathers in interviews.	A survey in Nigeria only asks mothers, ignoring that fathers often influence the decision either directly or indirectly.
Validity	The results really measure what the study claims to measure.	A mosquito net study measures malaria rates with blood tests, adjusts for weather, and rules out other causes.	A mosquito net study only asks people "do you feel healthier?" without checking actual malaria prevalence.
Reliability	The study would give similar results if repeated in the same way.	A survey tool is tested in two towns, gives consistent answers, and shows strong reliability statistics.	A survey tool gives very different results each time it is used, with no testing or explanation.
Cogency	The findings make sense from start to finish, and the study is honest about its limits.	A youth programme evaluation clearly links activities to outcomes, presents data, and admits long-term effects are uncertain.	A youth programme report claims "our activities worked" without showing data and ignores obvious limitations.

d. Make sure to capture findings systematically using a structured tool like a table or matrix. A good desk research will:

- Go beyond listing studies, instead synthesising insights across sources into a coherent picture of the situation and behaviours
- Include both situational and behavioural evidence, revealing how context and drivers interact
- Explicitly identify barriers and enablers
- Use credible, up-to-date, and context-specific sources, which are clearly cited and traceable
- Triangulate findings across multiple types of evidence (quantitative, qualitative, routine data, observational, feedback mechanisms)
- Clearly flag evidence gaps, contradictions, and uncertainties
- Frame insights in a way that is relevant and actionable for stakeholders and decision-makers

e. The format and structure of desk research can vary depending on the project's focus and objectives. Some reviews take a broad approach, summarising the overall context of the problem along with past interventions. Others are more targeted, exploring specific aspects such as common behavioural barriers, interventions that have already been tested, or the characteristics of the affected population. In general, a comprehensive desk research will include the following sections:

TABLE 4. SUGGESTED SECTIONS FOR A DESK RESEARCH

SECTION	DESCRIPTION
Cover Page	Includes title, date, author(s), organisation logos, and confidentiality note if required.
Executive Summary	Provides a concise overview of purpose, key findings, and main recommendations; designed to stand alone.
Introduction	Outlines the purpose, scope, and methodology of the desk research, including conceptual framework or definitions if relevant.
Context and Background	Presents sectoral, demographic, geographic, policy, and programmatic background; synthesises existing literature.
Key Findings	Core of the review, structured into sub-sections. For example: Main behavioural barriers, Key stakeholders related to the problem, Behavioural interventions that have been tested in the past
Gaps, Challenges, and Insights	Identifies limitations in data, evidence gaps, and conflicting findings.
Risk Analysis and Equity Lens	Highlight gaps related to gender, inclusion, human rights; surface risks to implementation.
Strategic Recommendations	Presents actionable and prioritized recommendations, tailored to decision-makers, implementers, and other audiences.
Conclusion	Summarises main takeaways, synthesises findings, and provides overall reflections.
References or Bibliography	Lists all sources cited and consulted during the desk research.
Annexes	Provides supplementary material such as detailed data tables, stakeholder lists, methods notes, or expanded case examples.

Some helpful templates and examples of desk research include UNICEF's desk research "Input Into Malaria Communication Strategy to Accelerate Malaria Elimination"; the joint UNICEF-NYU Ukraine desk research "The Power of Youth: Instruments for Effective Youth Participation"; and Save the Children's Desk Research Template Tool.

When applying a behavioural lens to the initial desk research

While not required at this stage, behavioural frameworks like UNICEF's **Behavioural Drivers Model (BDM)** can be a useful resource. The BDM is not intended as a tool for desk research or analysis, but rather as a conceptual guide that illustrates the multifactorial nature of behaviour. It reminds us that behaviours are rarely driven by a single factor, and that a full understanding requires attention to **psychological, social, and environmental** influences.

Keeping these domains in mind during the desk research process can help researchers notice patterns, surface blind spots, or better organize insights that may later inform diagnosis and design. This reflection is not about fitting data into a model, but about staying open to the layered realities that shape human behaviour. Consider the following prompts while reviewing existing evidence:

- What **psychological** factors — such as knowledge, beliefs, emotion, or self-efficacy — appear to influence this behaviour?
- What **social** influences — such as group norms, relational dynamics, or informal power structures — may be shaping decisions?
- What **environmental** conditions — such as availability of services, structural barriers, or institutional policies — create or constrain opportunities for action?

Learn more about the BDM: https://www.unicef.org/mena/reports/behavioural-drivers-model. *This model can also be used in conjunction with other frameworks, like COM-B.*

3. Determine whether primary research is needed to fill gaps

The final step in this phase is to assess whether existing evidence is strong enough to begin designing solutions or if further insights are needed. Additional research should only be conducted if it will meaningfully strengthen the intervention design process.

a. Review the desk research findings, and ask:

- Do we have enough of an understanding of the behavioural drivers, barriers, and context? For example, if the evidence identifies low clinic attendance but doesn't explain why - such as time constraints, negative provider experiences, or social norms — further inquiry may be needed.
- Is current evidence lacking any notable groups, or key voices from underserved or marginalized communities? For instance, if the literature focuses on urban caregivers but the intervention will be implemented in rural areas, key contextual differences may be overlooked.
- Are existing insights outdated, biased, or lacking local relevance? A study from five years ago, or one conducted in a different country or cultural setting, may not reflect the lived realities or current behaviours of the target population.
- Are there any major assumptions or unknowns that still need to be tested? For example, assuming caregivers forget appointments may lead to reminder-based interventions; however, if vaccine hesitancy or mistrust is the core issue, this solution may prove ineffective.

b. If key questions remain unanswered, conducting targeted research can help to fill in gaps in understanding. Large-scale studies are not always needed. Additional research could include, for example, engaging with a group that has yet to be consulted, or exploring how a social norm is reinforced, challenged, or maintained within a particular community context. Primary research should be focused and purposeful, avoiding the collection of irrelevant data, the overburdening of communities, or the delaying of the overall DEPTHS process.

c. After completing the desk research, return to the Define outputs to update them with any new insights. The Project Canvas, in particular, serves as a living document, anchoring the project's direction and informing all subsequent decisions. Updating the Canvas ensures consistency and alignment across the project.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

This initial desk research tool was not developed by the original project team. It is a recreated example based on real project data and context.

Conducting a desk research to understand existing data and evidence about the problem

To begin, the project team reviewed the key outputs generated in the Define phase (The Project Canvas, Problem and Outcome Statement, Behaviour Tree, etc) to determine their existing knowledge of the problem and any research gaps.

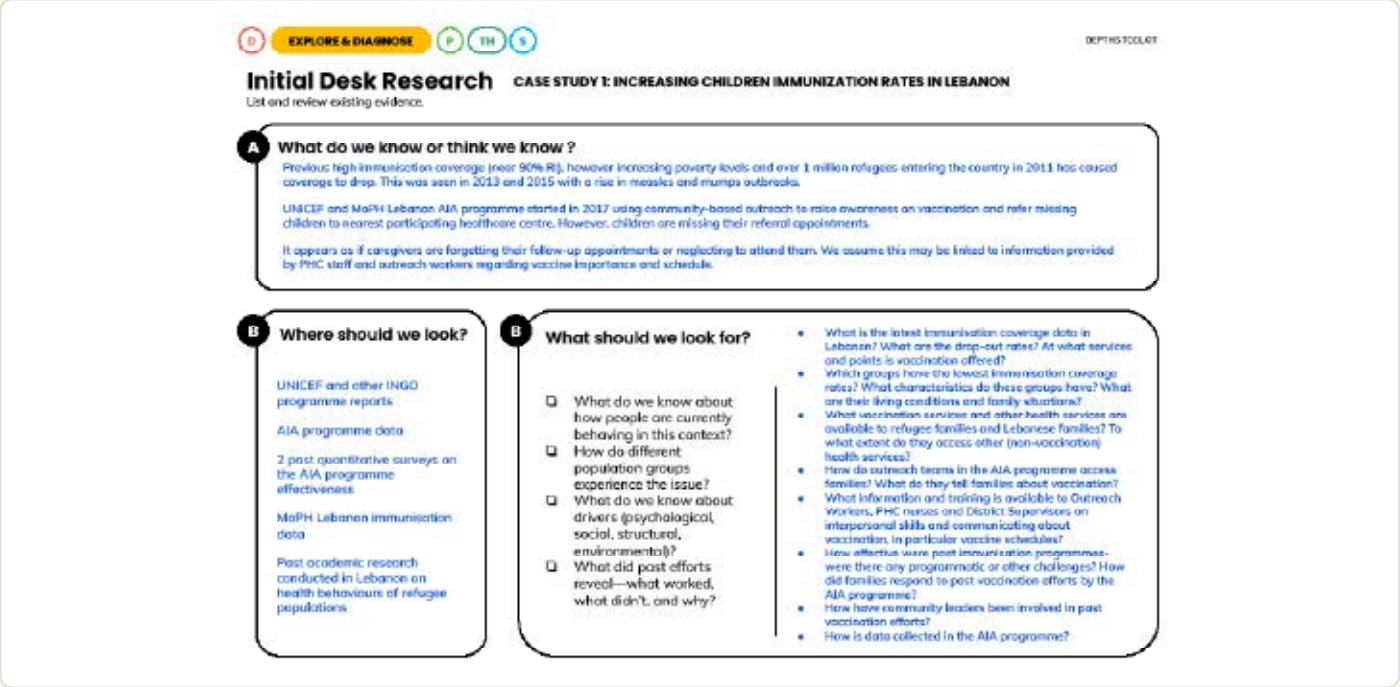
The team then gathered existing documentation, including programme reports, previous research and immunization coverage data within Lebanon, along with data from similar contexts (e.g. Jordan, Turkey, and Syria). Documents reviewed included:

- Past UNICEF and WHO immunization programme reports from Lebanon and similar contexts in Middle East and North Africa (MENA)
- Academic research on vaccine hesitancy and dropouts from low- and middle- income countries (LMICs)
- Immunization coverage data from Lebanon
- Qualitative and quantitative research data from past programmes
- Health Systems Data

The researchers reviewed and synthesised key information to improve their contextual understanding of the project's problem statement.

To ensure an efficient desk research process, the project team developed a plan to focus on the key areas and documents to explore and review.

Once the desk research was complete, the researchers organized key findings into a structured spreadsheet. This served as a central repository of relevant data, behavioural insights, and identified gaps. While the spreadsheet can be a final output in some cases, it more often supports a formal synthesis effort — such as a literature review, summary report, or insight synopsis. These products help to distill patterns, surface remaining questions, and inform the focus of subsequent primary research.



D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Initial Desk Research CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

List and review existing evidence.

A What do we know or think we know ?

Previous high immunisation coverage (near 90% RI), however increasing poverty levels and over 1 million refugees entering the country in 2011 has caused coverage to drop. This was seen in 2013 and 2015 with a rise in measles and mumps outbreaks.

UNICEF and MoPH Lebanon AIA programme started in 2017 using community-based outreach to raise awareness on vaccination and refer missing children to nearest participating healthcare centre. However, children are missing their referral appointments.

It appears as if caregivers are forgetting their follow-up appointments or neglecting to attend them. We assume this may be linked to information provided by PHC staff and outreach workers regarding vaccine importance and schedule.

B Where should we look?

UNICEF and other INGO programme reports

AIA programme data

2 past quantitative surveys on the AIA programme effectiveness

MoPH Lebanon immunisation data

Past academic research conducted in Lebanon on health behaviours of refugee populations

B What should we look for?

- ❑ What do we know about how people are currently behaving in this context?
- ❑ How do different population groups experience the issue?
- ❑ What do we know about drivers (psychological, social, structural, environmental)?
- ❑ What did past efforts reveal—what worked, what didn't, and why?

- What is the latest immunisation coverage data in Lebanon? What are the drop-out rates? At what services and points is vaccination offered?
- Which groups have the lowest immunisation coverage rates? What characteristics do these groups have? What are their living conditions and family situations?
- What vaccination services and other health services are available to refugee families and Lebanese families? To what extent do they access other (non-vaccination) health services?
- How do outreach teams in the AIA programme access families? What do they tell families about vaccination?
- What information and training is available to Outreach Workers, PHC nurses and District Supervisors on interpersonal skills and communicating about vaccination, in particular vaccine schedules?
- How effective were past immunisation programmes- were there any programmatic or other challenges? How did families respond to past vaccination efforts by the AIA programme?
- How have community leaders been involved in past vaccination efforts?
- How is data collected in the AIA programme?

STEP 2:
Develop primary research questions

Associated Tools:

Primary Research Objectives and Questions

In this step:

With the initial desk research complete, the next step is to define the specific behavioural questions that are in need of further exploration through primary research — using the Primary Research Questions worksheet.

This tool ensures that research questions are informed by behavioural science, allowing teams to diagnose the full range of drivers and barriers that shape behaviours.

Why it matters:

Clear and targeted research questions are essential to behaviourally informed inquiry.

Poorly defined or overly general research questions can result in data that is vague, redundant, or difficult to translate into action. In contrast, strong research questions sharpen the focus of inquiry, help to ensure that research instruments are well-aligned with the objectives, and enhance the rigour and relevance of findings.

Refining research questions at this stage achieves the following outcomes:

- **Ensures that research is aligned with the behavioural outcome** articulated during the *Define* phase
- **Focuses the investigation on priority behavioural determinants** at different levels, when questions are grounded in a behavioural model or framework
- **Avoids duplication of existing knowledge**, ensuring that primary data collection is focused on filling genuine gaps in evidence

To help structure this process, use the **COM-B**[2] model: a practical and widely used approach to diagnose the factors that or inhibit the behaviour of interest. Using a model like COM-B allows teams to systematically map why people do or don't perform behaviours.

The COM-B model is a framework that breaks behaviour down into three essential components: **Capability**, **Opportunity**, and **Motivation**, each with two sub-domains.

- **Capability** refers to the psychological and physical capacity to engage in the activity, including knowledge, skills, and mental faculties.
  - **Physical Capability** refers to physical skill, strength, stamina, or ability needed to perform the behaviour.
  - **Psychological Capability** refers to the mental or cognitive capacity to engage in the necessary thought processes, such as comprehension, reasoning, and memory.
- **Opportunity** encompasses the external conditions that make the behaviour possible, including environmental, socio-cultural, and structural factors.
  - **Social Opportunity** refers to the social cues, cultural norms, and interpersonal influences that shape behaviour.
  - **Physical Opportunity** refers to the environmental factors like resources, time, infrastructure, and access that enable behaviour.
- **Motivation** includes both reflective processes (such as beliefs and intentions) and automatic processes (such as emotional responses and habits) that drive behaviour.
  - **Reflective Motivation** refers to conscious planning and evaluation processes, such as intentions, beliefs, and identity.
  - **Automatic Motivation** refers to the unconscious processes such as emotions, habits, impulses, and reactions.

2 Michie, S., van Stralen, M.M. and West, R., 2011. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science, 6(1), p.42. https://doi.org/10.1186/1748-5908-6-42

Table 5 below outlines examples of behavioural drivers and barriers for each COM-B sub-domain.

TABLE 5. MAPPING OF BEHAVIOURAL DRIVERS AND BARRIERS IN COM-B FRAMEWORK

SUB DOMAINS	TYPES	DEFINITION
Psychological capability	Knowledge / understanding	How well someone understands the behaviour, service, or related information.
	Memory, attention, decision	How cognitive load, distractions, or competing priorities affect decisions.
	Behavioural regulation	How well someone can monitor and adjust their own behaviour.
Physical capability	Physical skills	The specific physical abilities or skills needed to carry out the behaviour.
	Strength, stamina, mobility	Energy, strength, or physical condition needed to act consistently.
	Health status	Health conditions that may enable or prevent action.
Physical opportunity	Time and scheduling	The ability to act within available time or schedules.
	Availability	Whether services, tools, or products are available when needed.
	Accessibility	Whether people can access services or places (e.g., distance, transport, hours).
	Structural barriers	Physical or systemic barriers that make action difficult or impossible.
Social opportunity	Norms and expectations	What is considered typical, acceptable, or expected in the social group.
	Social support	Support or encouragement provided by others.
	Influence of others	Direct influence from family, peers, leaders, or others in one's network.
	Roles and responsibilities	The roles people are expected to play, and how those shape their behaviour.
Reflective motivation	Beliefs about consequences	What someone believes will happen if they act (or don't).
	Self-efficacy	Whether someone believes they are capable of taking action.
	Intentions and goals	The goals or intentions people set for themselves.
	Identity	How behaviour aligns with a person's sense of self.
Automatic motivation	Emotions	Feelings that motivate or demotivate action.
	Habits	Routine actions performed without much thought.
	Impulses	Sudden urges that may override conscious decisions.
	Cognitive biases	Mental shortcuts or distortions that shape decisions.

Using a framework like COM-B offers several benefits in applied behavioural science projects, such as:

- **Holistic diagnosis**: Considering multiple drivers of behaviour rather than defaulting to a single explanation (e.g. lack of knowledge).
- **Evidence translation**: Serving as a bridge between data and action. By mapping qualitative and quantitative findings to COM-B components, teams can turn insight into targeted design decisions.
- **Strategic focus**: Prioritizing where interventions should focus. For instance, if motivation is high but physical opportunity is low, interventions may need to address structural barriers, not awareness.
- **Supports systems thinking**: Encouraging teams to explore how individual actions are shaped by context, norms, institutions, and infrastructure.
- **Flexible application**: Applicable at any point in the programme cycle: to shape formative research, identify entry points for behaviour change, refine prototypes, or explain why a pilot did or didn't work.

How to do it:

1. Craft research objectives that will guide the primary research process.

Any primary research will build on the initial desk research by addressing specific gaps that have been identified. Rather than listing broad themes or topics, define what must be understood in order to design context-specific, behaviourally informed, practical and promising interventions. Using the COM-B model to structure research objectives ensures a systematic exploration of important factors influencing the target behaviour.

- Review the updated Project Canvas and desk research findings to clarify what is already known, what is assumed but unverified, and what remains unknown, particularly regarding behavioural drivers, lived experiences, and system dynamics.
- Translate knowledge gaps into focused primary research objectives, writing them down in the Research Objectives and Questions worksheet. Avoid generic objectives — instead, focus on the behavioural mechanisms, contextual factors, or system dynamics that truly matter.
- Frame each objective around what still needs to be understood to influence the behaviour. Start with clear, purposeful verbs that signal the focus of the inquiry, such as:
    - *To understand…*
    - *To explore…*
    - *To identify…*
    - *To examine…*

  Each research objective should clearly signal which part of COM-B it addresses. Table 6 outlines the difference between strong and weak research objectives, using the example of routine immunization:

TABLE 6: FORMULATING STRONG VS. WEAK COM-B RESEARCH OBJECTIVES

COM-B DOMAIN	WEAK RESEARCH OBJECTIVE	WHY IT IS WEAK	STRONG RESEARCH OBJECTIVE
Capability	To explore parents' perceptions of vaccination	Broad, vague, and doesn't identify an actionable barrier.	To understand how misconceptions about vaccine safety influence parents' decisions to vaccinate their children.
Opportunity	To understand access to services	Lacks clarity on which aspect of access is being studied.	To examine how distance to health services, transportation costs, and clinic hours affect timely childhood immunization in rural areas.
Motivation	To explore caregivers' attitudes towards vaccination	"Attitudes" is too broad and doesn't identify specific motivational mechanisms.	To identify the beliefs, emotions, and social influences that shape caregivers' motivation to complete the full immunization schedule.

Keep these tips in mind when crafting strong research objectives:

- **Targeted primary research does not need to explore everything about the behaviour.** Instead, it should focus on the behavioural drivers and barriers that still need to be understood to inform effective intervention design.
- **Objectives should link to a design or strategy need.** Every objective should support a decision. Ask: Will answering this question help to design a better intervention? If not, refine the questioning.
- **Avoid restating the outcome.** Do not repeat the end goal (e.g., "to improve ANC attendance"). Instead, probe the underlying behavioural dynamics. An improved version would be: "To identify the socio-cultural and structural factors influencing first-trimester ANC attendance among newly pregnant adolescents."

2. Formulate primary research questions

Clear and purposeful research objectives are next translated into structured primary research *questions*. These questions shape data collection and guide the analysis process.

A research objective is a broad, action-oriented statement that outlines what the study seeks to achieve. In contrast, a research question defines exactly what the study will investigate. Research questions also guide what data to collect and who to engage. For example, if a research objective is "to understand how social norms influence adolescent girls' decisions to attend school during menstruation," related research questions can include:collect and who to engage. For example, if a research objective is "to understand how social norms influence adolescent girls' decisions to attend school during menstruation," related research questions can include:

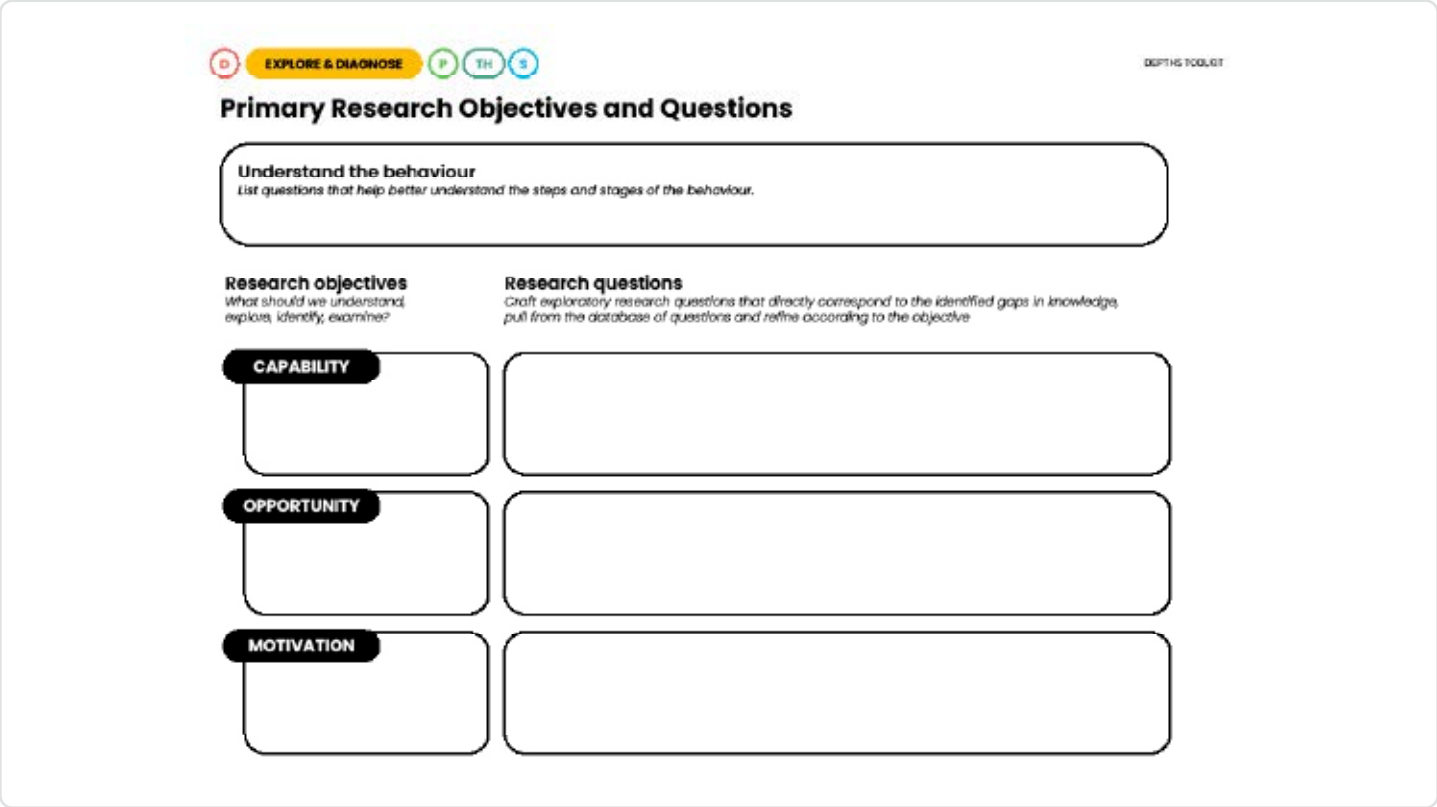
- What do peers and other girls in their community typically do when menstruating?
- How do peers, parents, and teachers influence girls' decisions to stay home or attend school?
- What forms of support or stigma do girls experience during menstruation, and how does this affect their behaviour?

A good research question meets the following criteria:

- **focused** on a particular issue (behaviour, belief, barrier or group)
- **answerable** through available research methods
- **yields insights that can inform actions** (decisions, designs and strategies)
- **shaped by an understanding** of what drives people's behaviour, often exploring the why or how behind specific actions or choices.

Here's how to formulate primary research questions:

a. First, fill in the "Understand the behaviour" section of the Primary Research Objectives and Questions worksheet. Before exploring behavioural drivers, it's clarify how the behaviour itself unfolds in practice. This means understanding what the behaviour looks like on the ground, when and


D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Primary Research Objectives and Questions

Understand the behaviour
List questions that help better understand the steps and stages of the behaviour.

Research objectives *What should we understand, explore, identify, examine?*	**Research questions** *Craft exploratory research questions that directly correspond to the identified gaps in knowledge, pull from the database of questions and refine according to the objective*
CAPABILITY	
OPPORTUNITY	
MOTIVATION	

where it happens, what steps are involved, and where things tend to break down. These descriptive questions ensure the research is grounded in the real-world experiences of those most affected, making later diagnosis of behavioural drivers more accurate and useful. Exploratory questions to understand the behaviour at hand can include, but are not limited to:

- When and where does the behaviour typically occur?
- What steps or decisions are involved?
- What are common drop-off points, or moments of friction?

For example, in the case of childhood immunization:

What are the steps caregivers take to get their child vaccinated — from first hearing about immunization, to deciding to go, to arriving at the clinic? At what age or stage do families typically begin or stop the immunization schedule, and what challenges emerge along the way?

b. Using the **COM-B model** helps to structure research questions and frame the inquiry Capability, Opportunity, and Motivation to influence a specific behaviour. Refer to the Research Questions Database to find prompts to help develop questions for each COM-B domain. Make sure to adapt them to the behaviour and populations of interest.

The following tables pair research objectives with strong vs. weak research questions, using an example of routine immunization:

COM-B DOMAIN

Capability

RESEARCH OBJECTIVE: To understand how perceptions or misconceptions about vaccine safety influence parents' decisions to vaccinate their children.

WEAK RESEARCH QUESTION	WHY IT'S WEAK	STRONG RESEARCH QUESTION
What do caregivers know about vaccines?	Vague — this could refer to any aspect of vaccines (e.g. cost, access, experience)	What safety beliefs and misconceptions lead caregivers of children under two in Kano (Nigeria) to start, delay, or skip scheduled vaccine doses?
Have caregivers received training or health talks on vaccines?	Focuses on format rather than content or beliefs, without addressing how or why misconceptions form.	Which information sources do caregivers in Kano (Nigeria) use when deciding whether to attend their child's next vaccination visit, and how does source credibility affect attendance?
Do caregivers know vaccines are safe?	Assumes yes or no answer, indicating binary and closed thinking.	To what extent do caregivers in Kano (Nigeria) believe routine vaccines are safe, and how does that belief affect starting, delaying, or skipping scheduled doses?

COM-B DOMAIN

Opportunity

RESEARCH OBJECTIVE: To examine how distance to health services, transportation costs, and clinic hours affect timely childhood immunization in rural areas.

WEAK RESEARCH QUESTION	WHY IT'S WEAK	STRONG RESEARCH QUESTION
Is it easy for people to get to the clinic?	Vague and subjective — "easy" varies across respondents.	How far and how long do caregivers in rural Kano (Nigeria) travel to vaccination sites, and how does travel burden affect timely attendance at vaccine appointments?
What is the state of rural health systems?	Too broad and not linked to the specific behaviour.	How do clinic opening hours, staff availability, and service organisation in rural Kano (Nigeria) affect completion of the full immunization schedule for children under two?
Do people want more clinics nearby?	Hypothetical and solution-focused, skipping the behavioural insight.	What location and scheduling constraints lead caregivers in rural Kano (Nigeria) to miss or reschedule vaccination appointments?

COM-B DOMAIN

Motivation

RESEARCH OBJECTIVE: To identify the beliefs, emotions, and social influences that shape caregivers' motivation to complete the full immunization schedule.

WEAK RESEARCH QUESTION	WHY IT'S WEAK	STRONG RESEARCH QUESTION
Do caregivers think vaccines are important?	Oversimplifies belief systems into a yes or no view and doesn't explore why those beliefs matter.	What do caregivers of children under two in Kano (Nigeria) believe about the importance and benefits of completing the full immunization schedule, and how do these beliefs affect completion?
Are caregivers scared of vaccines?	Leading and narrow-focused line of questioning.	Which emotions do caregivers in Kano (Nigeria) experience at key points in the vaccination journey, and how do these emotions affect return visits?
Who influences vaccination decisions?	Too vague — doesn't explore how or why the influence matters, or what kind of influence it is.	Who influences caregivers in Kano (Nigeria) to vaccinate or delay, and in what ways do spouses, grandparents, peers, and local leaders shape return for scheduled doses?

Write down the selected research questions in the Research Questions section of the Primary Research Objectives and Questions worksheet next to their corresponding objective. Each objective should be paired with a few specific, research questions that make it possible to investigate the issue.

Here are some tips for developing strong research questions:

- **Specify who the question concerns**, including key subgroups if relevant.
- **Specify the behaviour of interest**, as a single, observable action.
- Optionally **specify where the behaviour occurs**, as context can vary across settings.
- **Avoid vague, binary, or double-barreled questions**. For example, replace "Do people trust health workers?" with "What factors influence caregivers' trust in health workers during vaccination visits?"
- **Split broad or compound questions into clear parts**. For example, change "Why do people delay care and not follow up?" to two questions: "What causes delays in seeking care?" and "What prevents follow up after the first visit?"
- **Ensure the questions are decision-relevant**. Before finalizing a question, ask: Will the answer to this help us design more effective solutions? For example: What structural barriers do pregnant adolescents face when accessing antenatal care in informal urban settlements?

CASE STUDY

Increasing childhood vaccination uptake in Lebanon

The Primary Research Objectives and Questions were not developed by the original project team. They are recreated examples based on real project data and context.

Develop clear research objectives

In the previous step of this phase, the research team structured and conducted an initial desk research to assess what was already known about vaccine uptake in the country. Their initial observations highlighted a worrying drop in routine immunization (RI) coverage following the 2011 influx of refugees and rising poverty levels. Despite outreach efforts through the UNICEF-Ministry of Public Health's (MoPH) AIA programme, many caregivers were not following through on referral appointments, suggesting potential behavioural gaps that traditional programmatic approaches hadn't resolved.

Using this early review, the team surfaced a number of working assumptions: that caregivers may be forgetting vaccination appointments, that primary health care (PHC) staff and outreach workers may not be communicating clearly about vaccine schedules, and that daily stressors might be interfering with caregivers' ability to prioritize follow-up visits. These hypotheses were grounded in preliminary AIA programme reports, past quantitative surveys, and government immunization data, but they revealed important gaps that needed further primary research.

D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Primary Research Objectives and Questions

CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

Understand the behaviour

How do caregivers in Lebanon access vaccination? How do they learn about vaccination and the vaccine schedule? What steps do families take to access vaccination services?

	Research objectives	Research questions
CAPABILITY	To identify caregivers' knowledge and gaps in caregivers' understanding of the vaccination schedule, including timing, number of doses, and next steps after AIA outreach visits. To explore whether PHC staff and outreach workers have the knowledge and confidence to consistently communicate accurate, understandable vaccination information.	• What do caregivers, in particular from refugee and host community households, understand about the vaccination schedule (e.g., required doses, timing, and sequence)? To what extent do they understand the immunization calendar? • Do caregivers know where and when to access vaccination services after outreach visits? • Are there knowledge gaps or misunderstandings that cause caregivers to delay or skip appointments? • How do outreach workers and PHC staff communicate vaccine information to caregivers? Are they confident in delivering clear and correct advice? • What skills or support do PHC staff need to accurately track missed doses and plan follow-ups?
OPPORTUNITY	To understand the physical, logistical, and structural barriers (e.g., clinic accessibility, mobility constraints, service hours) that affect caregivers' ability to vaccinate. To examine how social norms, household dynamics, and community beliefs shape caregivers' decisions to vaccinate.	• What do caregivers think of the AIA programme vaccination experience? Are there any practical issues or challenges e.g. distance to facility, quality of service, time taken, clinic opening hours, that prevent caregivers from attending scheduled appointments? • What are the community structures and dynamics of refugee families within Lebanon? • What role do social or community norms play in impacting vaccination? What do caregivers believe other families are doing in relation to vaccination? Who are the key influencers, champions or community leaders who communities would listen to for information on vaccination and vaccine schedules? • How are decisions related to child immunisation made within these families? Any differences between women and men or age?
MOTIVATION	To explore the beliefs, emotions, and values that influence caregivers' perceived urgency or importance of completing all recommended vaccinations on time. To identify factors contributing to trust or mistrust in vaccination services and how these perceptions impact motivation to act on outreach or PHC guidance.	• How do caregivers weigh the importance of timely vaccination compared to other daily priorities or stressors? • What emotional responses (e.g., worry, pride, guilt, confusion) do caregivers experience in relation to vaccinating their children? • Are there beliefs or past experiences that shape caregivers' trust (or mistrust) in health providers or the vaccination process? • What motivates outreach workers and PHC staff to deliver a positive vaccination service and/or go beyond their routine duties (e.g., follow up visits, repeated messaging)?

The team used the COM-B model to guide the development of their research objectives. This model gave them a structured way to think about the various factors influencing caregivers' behaviour. Under Capability, they sought to explore knowledge gaps; for example, whether caregivers understood when and where to return for vaccinations. Under Opportunity, they wanted to understand how clinic hours, transportation, gender norms, and peer influence affected follow-through. And under Motivation, they aimed to examine emotions, trust, and competing priorities that could affect caregivers' decisions to act on referral guidance.

Develop primary research questions

With the COM-B framework in hand, the team translated these focus areas into precise primary research questions. For example, they asked: "How do caregivers weigh the importance of timely vaccination compared to other daily priorities or stressors?" and "What do caregivers know about where and when to access vaccination services after outreach visits?" These questions were designed not only to diagnose behavioural barriers, but also to help the team uncover more nuanced insights around household decision-making and provider-caregiver dynamics.

The resulting research questions served as a foundation for designing tools to capture rich, contextual evidence from Lebanese and Syrian caregivers, health workers, and outreach staff. The questions also helped highlight what wasn't yet fully understood: such as the emotional trade-offs caregivers make when deciding to delay or skip vaccination or the role of informal community reminders.

STEP 3:
Plan the primary research

Associated Tool:

Primary Research Plan

[External] Research Protocol Template

In this step:

Step 3 translates the research questions into a practical plan. Each question is linked to the information needed, the people best placed to provide it, and the setting where honest, accurate answers are most likely (for example, at a clinic during busy hours to understand waiting times, or at home to explore decision making in the family).

Methods are chosen to fit the question: interviews for how and why, short surveys for how many, or observation when actions are easier to see than describe. The plan also sets out who will be included, how participants will be invited fairly, and the simple tools to keep conversations consistent (such as interview or discussion guides). It explains how notes, recordings, or forms will be used, and how privacy will be protected.

Where topics are complex or involve sensitive groups, a fuller protocol is prepared and ethics approval is required to ensure safety and responsibility. By the end of the step, the plan clearly links questions to methods, participants, and places; the data collection tools are ready to use; and, when required, an approved protocol and consent materials are in place. A brief pilot and an early team debrief help confirm that the plan is workable and that the information collected will answer the original questions and support programme decisions.

Why it matters:

It's essential to gather the right insights from the right people in ways that are ethical, actionable, and connected to the behaviour we aim to influence. Without a clear plan, there is a risk of collecting data that is too broad, shallow, or disconnected from the behaviour prioritized in the Define phase.

A strong plan also helps to appropriately allocate time and resources. It clarifies roles, identifies where local support is needed, and allows for adaptations to research methods, tools, and materials based on language, literacy, or cultural context. Most importantly, it ensures that teams are not only operationally prepared but also aligned on learning objectives.

How to do it

1. Identify who we need to hear from and how best to engage them

This includes not only the target population, but also members of the broader community and other knowledge sources. Each research question should be matched with the people who are best placed to answer it. Make note of whether any of these groups may be considered vulnerable or have less agency to determine whether they are comfortable participating.

Consider including:

- Primary participants with lived experience of the behaviour (e.g., caregivers in areas with low vaccination rates).
- Influencers who shape decisions (e.g., family members, health workers, peers).
- Context experts who understand systems or structures influencing the behaviour (e.g., local NGOs, government partners).

2. Identify research methods

To uncover why individuals may not engage in a target behaviour, carefully consider the methodological approach.

A range of research methods exist to understand the underlying causes of a given behaviour, and each method has advantages and limitations. Whenever possible, it's worth advocating for a mixed-methods approach, drawing on both qualitative and quantitative techniques for a richer, more holistic understanding:

- Qualitative methods (such as interviews, focus groups, and observations) help to uncover the why behind behaviours, exploring barriers and drivers.
- Quantitative methods (such as surveys) help to measure the what, identifying patterns, frequencies, and relationships across a larger population.

In many settings, resource constraints — whether related to time, budget, personnel, or geographic access — may limit the scope of methods employed. In such instances, it becomes essential to strategically select the approaches that are not only methodologically sound, but also contextually appropriate and logistically feasible. To select the most appropriate methods:

a. Clarify constraints and operational realities. Consider:

- What is the current capacity of the team, both in terms of expertise and availability?
- Will the research be conducted in person, or must it be implemented remotely?
- Will additional support be required? For example, interpreters, translators, or trusted local facilitators familiar with community dynamics.
- Are there temporal constraints, such as seasonal access, safety concerns, or participant availability, that might affect when and how data can be collected?
- Are there financial constraints that may affect the choice of methods, sample size, or field logistics?

b. Different methods serve different purposes. Table 8 below summarizes some of the commonly used behavioural research methods, with guidance on when each may be most useful. Rather than assigning one method per question, consider how a combination of methods can help triangulate findings, offset limitations, and deepen understanding.

TABLE 8. QUALITATIVE RESEARCH METHODS

METHOD	USEFUL FOR	STRENGTHS	LIMITATIONS
In-depth semi-structured interviews One-on-one conversations guided by a set of questions to explore experiences and motivations	Exploring individual beliefs, motivations, and sensitive topics in depth.	Flexible format allows participants to share personal experiences in their own words, fostering trust and openness.	Less scalable as interviews take more time, rely heavily on the interviewer's skill and neutrality, and typically yield insights that are not easily generalizable across populations.
Focus Group Discussions Facilitated discussions with a small group to uncover norms, perceptions, and shared experiences.	Understanding group norms, social dynamics, and shared attitudes.	Efficiently surfaces shared views and social influences, revealing how people respond to others' opinions and expectations.	Group settings may discourage open discussion of sensitive topics or create risks for participants. Dominant participants can steer the conversation, making it harder to capture diverse perspectives.
Field Observations Watching how behaviours and interactions unfold in natural settings.	Capturing real-world context and observable behaviours.	Provides direct insight into routines, environment, and interpersonal dynamics, including barriers people may not articulate.	If not well introduced, this type of data can feel intrusive — also more open to interpretation, and requires careful coding, note-taking, and often triangulation with other methods to make sense of patterns.

METHOD	USEFUL FOR	STRENGTHS	LIMITATIONS
Mystery User A trained person/researcher simulates a real user experience to assess a process and identify hidden barriers.	Testing service quality or frontline behaviour without observer bias.	Reveals breakdowns or user experience gaps that staff may not notice or report. It is especially useful in service or referral systems.	Requires well-trained assessors and may raise ethical concerns if participation is not disclosed or consented to.
Social network mapping *To identify influencers, social expectations, and approval/ disapproval dynamics around behaviours.*	Understanding who influences behaviour, how information and norms spread, and where social pressure or support comes from.	Reveals how behaviours are shaped by relationships, not just individual beliefs. Helps to identify people or groups who hold influence over decisions, making it easier to design targeted messages or leverage trusted messengers.	Accurately mapping networks takes time and careful facilitation, especially in communities where relationships are sensitive or hidden. People may hesitate to name influencers honestly, or may overlook informal connections that play a key role in shaping behaviour.

TABLE 8B. QUANTITATIVE RESEARCH METHODS

METHOD	USEFUL FOR	STRENGTHS	LIMITATIONS
Surveys *Structured questionnaires that gather quantitative data from a larger population.*	Measuring prevalence, attitudes, or self-reported behaviours across a population.	Allows data to be collected from large numbers of people in a consistent way, making it easier to compare data across locations or groups. Responses can be quickly summarised using simple statistical analysis, especially when using digital tools or pre-coded questions.	Surveys rely on what people say they do, not what they actually do. Responses can be influenced by what participants think is expected or socially acceptable (social desirability bias), and if questions are poorly phrased, confusing, or culturally misaligned, they can lead to misleading or incomplete results.
Service or administrative records *Routinely collected data from health systems, education records, registration logs, or other service delivery platforms.*	Understanding actual service uptake, identifying drop-off points, and tracking trends over time.	Uses real-world data to reveal actual behaviour, not only self-reports, often already available, and cost-effective for longitudinal trends.	Data may be incomplete, outdated, or inconsistently recorded — may not include reasons why behaviours occur.

Use the following considerations to help assess which methods are likely to work best for the context, population, and topic of interest:

- **Research goals and the nature of inquiry**
  - Is the goal to understand why something is happening (exploratory), or how often it occurs (descriptive)?
  - Is the aim to uncover deep motivations, beliefs, and decision-making processes — best suited to qualitative methods — or to identify broad patterns and trends across a population, which may require quantitative approaches?
  - Will behaviours be observed in context, or will teams rely primarily on self-reported accounts of beliefs and experiences?
- **Participant characteristics**
  - Are participants likely to feel safe and comfortable speaking openly about the topic? For sensitive issues, such as gender-based violence or reproductive health, individual interviews may be preferable to group settings.
  - Are there language, literacy, or accessibility considerations that could affect participation?
  - Would participants be more at ease in group discussions, or would they engage more freely in one-on-one interactions?
- **Cultural and contextual relevance**
  - Are there locally resonant methods that may yield deeper insights? For instance, participatory storytelling, transect walks, or visual mapping exercises may feel more natural and engaging than structured interviews in some communities.
  - Could certain methods inadvertently reinforce existing power dynamics (e.g., formal interviews conducted by government officials or outsiders)?
- **Sensitivity of the topic**
  - How emotionally or socially sensitive is the topic at hand? Topics involving stigma, trauma, or personal loss often require private, trauma-informed approaches.
  - Does disclosure carry risks for participants, and if so, how can protective environments be created?
- **Need for comparability and standardization**
  - Is there a requirement to generate standardized or comparable data across time, populations, or geographies? In such cases, validated tools such as the BeSD survey on vaccine hesitancy may be appropriate.
  - Alternatively, is open-ended, exploratory data required to identify new insights, surface unanticipated dynamics, or refine behavioural hypotheses?
  - It may be helpful to consult a research specialist to select the right mix of methods and ensure the study design is both feasible and rigorous. UNICEF has different teams that could provide support, such as the BIRD Lab and UNICEF Evaluation.

Bridging the gap between reported and actual behaviour

Data collection often relies on what individuals say — namely, their beliefs, experiences, and self-reported behaviours, but it is critical to recognize the limitations of this. Behavioural science reminds us of two critical truths: what people say is not always what they do, and human memory is fallible and often biased.

Consider a relatable example from everyday life: how much movement we get on a daily basis. A meta-analysis[3] comparing self-report surveys of sedentary time with device-measured data (like accelerometers) found that people underestimated sitting time by an average of ~1.7 hours per day, even when asked in simplified formats. The largest gaps appeared when surveys relied on single-item questions such as "How many hours do you sit on a typical day?", showing how simple self-reports can systematically miss the mark. More detailed formats like time-use diaries or multi-item questionnaires reduced but did not eliminate the bias. These limitations do not imply that participants are intentionally deceptive, but cognitive processes, emotional states, and social dynamics influence recall, interpretation, and communication.

As researchers and practitioners, it is necessary to approach self-reported data with thoughtful rigour, designing research to both respect participants' perspectives and compensate for these inherent constraints. Common challenges associated with self-reported data include:

- **Social desirability bias**. Participants may offer responses they believe are expected, appropriate, or socially acceptable, particularly in group settings or when discussing sensitive topics.
- **Memory distortion**. Individuals often recall generalized routines or scripted behaviours rather than specific instances. As memory tends to compress repeated experiences into familiar narratives, important variations or context-specific nuances may be lost.
- **Vague or abstract responses**. Without deliberate prompting, participants may default to broad or conceptual answers, reflecting general attitudes rather than concrete behaviour. While such responses can offer valuable perspective, they may have limited utility when seeking to design interventions grounded in specific behavioural drivers.

By recognizing these patterns, it is possible to effectively tailor research instruments and facilitation techniques, ensuring that the data collected is not only meaningful and respectful but also behaviourally credible and practically useful. This can be achieved by:

- Strengthen inquiry through thoughtful question design that elicits richer and more reliable data:
  - Anchor responses in time: Reference specific timeframes (e.g., "In the past week…" or "Yesterday…") to aid memory and prompt more accurate recall.
  - Specify context and behaviour: Ground questions in concrete scenarios and observable actions. For example, "What did you do at the clinic?" is more illuminating than "Do you usually attend check-ups?"
  - Prioritize first-hand experiences: Frame questions around what the respondent personally did, rather than what they believe others did.
  - Follow with emotional and motivational probes: Once a behaviour is identified, explore underlying drivers such as feelings, beliefs, or intentions.
  - Avoid leading language: Pose questions in a neutral way rather than planting responses. For example, "What have you heard families say about vaccination" can yield more candid responses than asking "Do most families disapprove of vaccination?"
- Adapting approaches for sensitive or traumatic topics. When exploring highly personal or potentially distressing issues, it is critical to create a supportive environment and apply trauma-informed techniques:
  - Avoid direct disclosure: Participants may feel more comfortable engaging through hypothetical scenarios or third-person narratives.

3 Prince, S.A., Cardilli, L., Reed, J.L. et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 17, 31 (2020). https://doi.org/10.1186/s12966-020-00938-3

- Use vignettes or visual stimuli: These tools can help depersonalize the conversation, reduce the emotional burden, and spark thoughtful reflection.
- Ensure facilitator preparedness: Those conducting the research should be trained to foster a space that is respectful, empathetic, and free of judgment — particularly when discussing stigmatized or vulnerable experiences.

- Complementing self-reported data with other methods. To further strengthen the behavioural validity of our insights, we can triangulate self-reported data with additional approaches:
  - Direct observation: Observing behaviours within their natural context can reveal patterns or inconsistencies not easily captured through verbal reporting.
  - Participant journaling: Encouraging respondents to record their experiences — via written logs, voice notes, or digital diaries — can yield more immediate and authentic reflections.
  - Repeated engagement: Where feasible, brief follow-ups via SMS, phone calls, or subsequent visits can help capture changes over time and verify earlier responses.

By designing research with these strategies in mind, both the ethical integrity and analytical strength of inquiry is enhanced.

3. Build the research plan

Using the Primary Research Plan worksheet, create an actionable research plan that clearly articulates what is intended to be investigated, how the research team will approach it, and why it matters. It is worth mentioning that for more complex, resource-intensive, and/or ethically sensitive research, it is highly recommended to develop a full research protocol.

The research plan serves as both a planning tool or a foundational document to develop a more in-depth research protocol that can be later used for ethical review.

Below, there are a set of guiding questions to decide whether a research plan is sufficient, or a full research protocol is needed. If the answer is "Yes" to any of these questions, developing a research protocol is highly recommended.If the answer is "Maybe" to any of these questions, a research plan might be sufficient, but it is strongly encouraged to reach out to the organization's ethics team for further discussion. Finally, if the answer is "No" to all of these questions, a research plan might be sufficient. For additional support, it is recommended to reach out to specific evaluation teams within your organization or external research partners such as academic institutions or research firms. In UNICEF, reaching out to the BIRD Lab can be a good starting point.

TABLE 9. DECIDING BETWEEN A RESEARCH PROTOCOL OR PLAN

QUESTION	ANSWER	FURTHER GUIDANCE
Will ethics review or institutional approval likely be required? *(For further guidance on how to decide whether ethics review is needed, consult 'Sub-step 5: Ethics approval')*	☐ Yes ☐ No ☐ Maybe	Research protocols are often required by an ethics board; if ethics review and approval is expected, it is recommended to develop a research protocol.
Will there be data collection from or about vulnerable populations (e.g., children, refugees)?	☐ Yes ☐ No ☐ Maybe	Collecting data from vulnerable populations requires ethical approval. Some exceptions can be made if (a) the organization routinely collects data from vulnerable populations, (b) the research is part of its regular and existing processes and programmes, and (c) if the project is not going to be part of an academic publication.
Will the research involve collecting personal, sensitive, or health-related information?	☐ Yes ☐ No ☐ Maybe	Collecting sensitive data and information requires ethical approval. It is important to highlight that all sensitive information will require to be anonymized and have informed consent from participants.
Will the research involve audio, video, or image recordings of participants?	☐ Yes ☐ No ☐ Maybe	Audiovisual recordings are personally identifiable data. Collection of personally identifiable data for research generally requires ethical review.

(Table continues on following page)

QUESTION	ANSWER	FURTHER GUIDANCE
Will the research findings be submitted for academic publication or formal dissemination?	☐ Yes ☐ No ☐ Maybe	If the primary aim of the academic article is to present research findings, a formal research protocol will likely be required. However, if the research is formative and the article focuses more on describing an intervention or proposed solution, a research protocol may not be necessary — though it is still recommended.

To build a primary research plan or research protocol (if needed), make sure to include:

a. **Research objectives**: Return to the research objectives developed using the COM-B model in Step 2, and reflect on what to explore.

b. **Research questions:** Return to the research questions developed in Step 2 using the COM-B model. These questions translate each objective into focused lines of inquiry. Paired with objectives, these questions form the foundation of the research plan, guiding the selection of methods, the type of data to collect, and the approach to analysis. It is important to note that each objective may have multiple research questions.

c. **Research participants:** Identify the key population groups whose perspectives are essential to answering the established research questions. This includes not only the primary group whose behaviour is in focus, like caregivers, but also those who influence or enable the behaviour, such as family members, health workers, or community leaders. Revisit the Stakeholder and Audience Map worksheet from the Define phase. It is also important to note that the information gathered about the target population will be used as input to develop the Behavioural Profile tool in Step 5 of this phase (Analyse and Synthesise Findings). Be specific about:

   - Who to engage
   - How many participants from each group
   - What variation to capture (e.g., by gender, age, location, or social status)

   Documentation of meaningful consent is required when engaging participants in research activities. Meaningful consent entails that participants are informed of the purpose of the activity, how it will be conducted, how their data will be used, and any potential risks. For more about informed consent see childethics.com.

   Apply an intersectional lens to understand how different aspects of identity, such as gender, age, disability, or refugee status, shape people's experiences. Consider who may be excluded or silenced in typical research modes, and take actionable steps to include them. For example, when studying childhood vaccination, refrain from assuming that mothers are the only decision-makers or clinic-goers: fathers and grandparents may also play a key role.

   To uncover these dynamics, ask:

   - Who does what?
   - Who has what?
   - Who decides?

d. **Sampling**: The next step is determining how to identify and engage these individuals. Sampling is the process of selecting a subset of individuals from a broader population to participate in the research. The quality, relevance, and inclusiveness of findings are shaped by the sampling decisions made at this stage. There are a range of sampling strategies, each with distinct advantages and limitations, depending on the research context, objectives, and constraints. It is important to align the sampling strategy with the type of research that will be conducted.

   The following table summarizes some of the most commonly used sampling strategies, what they are most suitable for, when to use them, and their strengths and limitations.

TABLE 10. SAMPLING STRATEGIES

SAMPLING STRATEGY	MOST SUITED FOR	STRENGTHS	LIMITATIONS	WHEN TO USE
Simple random sampling is like drawing names from a hat. Every person in the group has an equal chance of being picked. For instance, if there's a list of all parents in a district, selecting 100 of them at random to ask about their children's vaccination status uses this method.	**Quantitative research**. Ideal for surveys or experiments needing statistical inference. Rare in qualitative research.	This method is fair and unbiased because everyone has the same chance of being selected. It helps to create results that represent the whole population.	This strategy requires a complete and accurate list of the population, which can be hard to attain in some areas.	Use when a complete list of the population is available, and the goal is to get a representative view of the whole group.
Systematic sampling selects every n^{th} person from a list, after starting at a random point. For example, export the clinic's eligible caregiver list (n=600) to a spreadsheet. Assign each record a random number and sort by that number. Pick a random start between 1 and 20, then select every 20th record to recruit n≈30.	**Quantitative research**. Easier than random sampling, still supports generalization. Can be used in qualitative research but rarely preferred.	This method is simpler than random sampling and still reduces bias, if the list is ordered randomly.	Similar to simple random sampling, this strategy also requires a complete and accurate list of the population.	Use when there is a well-organized list, and a simpler alternative to full random sampling is needed.
Cluster sampling is employed when it's difficult to reach every individual. The population is divided into groups (e.g., neighbourhoods or health districts) and a few of these groups are randomly chosen. Everyone within the chosen groups is then included. For instance, selecting three districts and interviewing all caregivers located there about vaccines.	**Mainly quantitative research**. Useful for large, geographically spread populations in surveys. Can work in qualitative research when entire settings (e.g., clinics or schools) are studied.	This method saves time and resources, especially in large or spread-out areas.	This method may give inaccurate results if the selected groups are very different from others that were not chosen.	Use when studying large or spread-out populations, where it's impractical to reach everyone individually.

(Table continues on following page)

SAMPLING STRATEGY	MOST SUITED FOR	STRENGTHS	LIMITATIONS	WHEN TO USE
Proportionate stratified sampling splits people into smaller groups (strata) like age, gender, or location, and selects participants in the same proportion as they appear in the whole population. For example, if 60% of caregivers live in urban areas and 40% in rural ones, the sample will reflect those proportions.	**Both**. Great for comparing groups quantitatively; may be used in qualitative research if diversity across strata is a goal.	This strategy ensures that key groups are represented fairly, improving the accuracy of the results.	This strategy requires detailed population information to divide people into the right groups.	Use when it's important to reflect the natural make-up of different groups within the population.
Disproportionate stratified sampling also splits people into strata, but it intentionally selects more people from certain groups (usually smaller or more important ones). For instance, it may focus more on rural caregivers in areas with low vaccination rates, even if they're a small part of the population.	**Both**. Great for comparing groups quantitatively; may be used in qualitative research if diversity across strata is a goal.	This method allows researchers to focus on important but small groups, helping uncover detailed insights.	This method can create bias in the overall results and could require adjustments to make the sample more balanced.	Use when certain subgroups are small but important to the research and need further representation in the sample.
Convenience sampling picks people who are easiest to reach and willing to participate. For example, interviewing caregivers who visit a vaccination clinic during a study period.	**Qualitative research**. Often used in early-stage or exploratory studies. Not ideal for inference, but extremely practical.	This method is quick, low-cost, and easy to carry out, making it useful when time or access is limited.	This method may not give a balanced view and results often cannot be applied to the general population.	Use in quick, early-stage, or low-resource research where easy access to participants is needed.
Judgmental purposive sampling selects people who are especially knowledgeable or relevant to the topic, based on the researcher's expertise. For instance, choosing experienced health workers in high-risk areas to discuss vaccine hesitancy.	**Qualitative research.** Classic in-depth qualitative method—targets key informants with relevant insights.	This method helps to gather rich, relevant insights from people with direct knowledge or experience.	This strategy may reflect the project team's personal bias in choosing participants.	Use when specific, knowledgeable people are best suited to answer the research questions.

(Table continues on following page)

SAMPLING STRATEGY	MOST SUITED FOR	STRENGTHS	LIMITATIONS	WHEN TO USE
Quota purposive sampling ensures a certain number of people from specific categories are included. For example, selecting an equal number of mothers and fathers to study gender differences in vaccine decisions.	**Qualitative research.** Ensures minimum variation (e.g., gender, age); common in focus groups or small-scale interviews.	This method guarantees that certain types of people are included, even without a full population list.	This method may still introduce bias if individuals that are easiest to find are not representative within the broader group.	Use when it's necessary to ensure a specific mix of participant types, without needing a full list of the population.
Snowball sampling starts with a few participants who then refer others. This method is helpful for finding people who are difficult to reach. For example, beginning with one parent who is hesitant about vaccines, and asking them to refer others they know.	**Mainly qualitative research**. Best for hidden or hard-to-reach populations. Can also work in mixed methods.	This method helps find and include people who are otherwise difficult to identify or reach.	This method can lead to biased results, as individuals often refer others who are similar to themselves.	Use when studying hard-to-reach populations that are connected through social networks.

The choice of sampling strategy should be guided by the nature of the research questions, available resources, the sensitivity of the subject matter, and the accessibility of different population groups. In many cases, a mixed approach may be most appropriate — for example, using purposive sampling to ensure the inclusion of key subgroups, alongside snowball sampling, to access individuals who may not be reached through conventional recruitment channels.

e. **Research methods and tools:** This section should describe the chosen data collection approaches and the reasoning behind them. Each method should be selected based on its relevance, feasibility, and fit with the context and participants. See Table 8 for guidance on selecting an appropriate research method. Clearly document methodological choices to strengthen the credibility of the research.

f. **Data collection materials:** It's important to prepare the right materials for effective and ethical data collection. These materials should be tailored to the selected method and adapted to the context, accounting for factors like language, literacy, cultural norms, and accessibility. Table 11 on the following page offers a summary of common materials needed for different research methods:

TABLE 11. DATA COLLECTION MATERIALS BY METHOD TYPE

METHOD	MATERIALS NEEDED
In-depth interviews (IDI)	☐ Interview guide aligned with research questions ☐ Audio recorder or transcription tool ☐ Note-taking template ☐ Informed consent forms
Focus group discussions (FGD)	☐ FGD discussion guide ☐ Visual prompts or handouts, if needed ☐ Flipchart/markers, if needed ☐ Audio recorder or transcription tool ☐ Note-taking template ☐ Informed consent forms
Field observations	☐ Structured observation checklist or open-ended template ☐ Contextual map or site layout, if needed ☐ Discreet note-taking materials ☐ Observer script (if applicable) ☐ Consent or ethical protocol (if disclosed)
Mystery user	☐ Mystery user script/scenario ☐ Evaluation checklist ☐ Debrief form ☐ Consent and ethical protocol
Surveys (digital or paper)	☐ Survey instrument (translated/adapted as needed) ☐ Data collection platform (e.g., KoboToolbox, ODK) ☐ Tablets or phones (if digital) ☐ Printed survey questionnaire (if paper) ☐ Enumerator script ☐ Informed consent forms
Social network mapping	☐ Mapping tool or template ☐ Prompt questions to identify influencers and relationships ☐ Drawing supplies and large paper (or digital whiteboard, if interactive)
Service or administrative records	☐ Formal agreement or written permission from the data owner (e.g. Ministry of Health, school administrator) to use the data ☐ Data dictionary or metadata (if available) to understand how variables are defined and recorded

g. **Timeline**: Develop a realistic, structured timeline to guide planning, data collection, and preliminary analysis. This should outline the duration and sequencing of activities, while accounting for logistical needs, potential delays, national holidays, and other context-specific constraints. The timeline should include:

- Finalization of research tools
- Time for obtaining all necessary field work approvals
- Ethical review (if required)
- Enumerator training
- Pilot testing
- Data collection periods (by location or target group)
- Time for transcription, translation, and data cleaning
- Preliminary and final analysis
- Check-ins, debriefs, or validation sessions with stakeholders
- Buffer periods and anticipated disruptions

A well-structured timeline serves as a shared reference point and encourages alignment across teams and stakeholders. The example below offers a basic template teams can adapt to fit the scope and complexity of their research:

Activity	Status	W1	W2	W3	W4	W5	W6	...	Lead/Owner
Finalize research tools	Done								Research manager
Obtain field work approvals	In progress								Field coordinator
Enumerator training	Not started								Research manager
....									

h. **Roles and responsibilities**: Delineate clear roles and responsibilities for each phase of the research process to ensure efficient coordination and accountability. This includes identifying team members responsible for data collection activities, as well as those managing logistics, documentation, analysis, and stakeholder engagement.

For example, a locally based facilitator may be tasked with coordinating field-level implementation and participant engagement, while the central research team oversees tool refinement, quality assurance, and data synthesis. Where external partners or consultants are involved, their contributions should also be clearly defined.

It is also helpful to assign one focal person responsible for quality assurance and troubleshooting across all data collection sites. This role serves as a key link between field teams and the central research team.

i. **Ethical Considerations:** The project team is responsible for upholding ethics principles for all research involving people, regardless of whether independent ethics review is mandated. In this section, outline how the proposed research will reflect ethics principles for evidence activities. For UNICEF, all activities should be screened through the lens of the "best interests of the child", and specifically reflect principles of benefit, do no harm, respect, fairness, accountability and integrity. This should include undertaking a preliminary risk assessment and developing a risk mitigation plan to address each identified risk, as well as consideration of standards such as how privacy will be maintained, how to ensure informed consent, and what measures may be required to safeguard participants.

4. Build research tools

With the research plan or protocol in place, the next step is to build the specific tools and materials needed to gather high-quality data. These include instruments such as interview guides, focus group discussion guides, surveys, and observation templates, each designed to capture different types of behavioural insights.

a. Start by selecting which tool need to be developed based on the selected research methods. See Table 8 for guidance on selecting a research method.

Table 12 summarises commonly used primary research tools, typical use cases, relevance to behavioural research, and relevant resources that offer inspiration or detailed guidance for developing similar tools. This is a non-exhaustive list; these tools are intended as starting points and should be adapted to reflect the specific behaviour, context, and population of interest.

TABLE 12: RESEARCH TOOLS

TOOL	RELEVANT RESOURCES
Survey	• WHO/UNICEF (2022), Behavioural and Social Drivers of Vaccination (BeSD) (incl. Childhood Vaccination Survey for Caregivers) • UNICEF (2018), Measuring Social and Behavioural Drivers of Child Protection Issues: Guidance Tool (incl. FGM/C and Child Marriage Standardized Surveys) • J-PAL, Survey Design
In-depth interview guide	• Pathfinder International (2001), Conducting In-Depth Interviews: A Guide for Designing and Conducting In-Depth Interviews for Evaluation Input • The Behavioural Insights Team (2022), Explore Field Guide: A practical tool to map and unpack behaviour
Focus group discussion guide	• Nyumba et al. (2018), The Use of Focus Group Discussion Methodology: Insights from Two Decades of Application in Conservation • The Behavioural Insights Team (2022), Explore Field Guide: A practical tool to map and unpack behaviour
Observation checklist	• The Behavioural Insights Team (2022), Explore Field Guide: A practical tool to map and unpack behaviour
Mystery user guide	• Ipsos (2020). Mystery Shopping: Seven Steps to Designing a Better Mystery Shopping Programme
Participatory activities *for use in IDIs and FGDs — for example, but not limited to: Cards sort, vignettes, day in the life, choices map, mind and body mapping, gender boxes, social network mapping*	• UNICEF (2019), Everybody Wants to Belong: A Guide to Tackling and Leveraging Social Norms in Behaviour Change Programming • UNICEF (2020), Participatory Research Toolkit for Social Norms Measurement • UNICEF/First Hand, Rapid inquiry in the HCD4health toolkit

b. Next, design the tools. Pay careful attention to how instruments are structured and how questions are framed, in order to elicit relevant, honest, and useful responses. Keep these suggestions in mind:

- **Structure questions around observable behaviours, decisions, and experiences**. For example, replace "Do you know how to keep your baby healthy?" with "Can you walk me through what you usually do to keep your baby healthy?".
- **Add questions that explore participants' lives** beyond the immediate behaviour of interest. Understanding daily routines, priorities, challenges, aspirations, and social relationships can reveal contextual factors that shape behaviour and will inform the development of Behavioural Profiles '(see' Step 5: Analyse and synthesise findings').
- **Use culturally resonant and accessible language**. Replace technical terms like "complementary feeding" with more colloquial language such as "Which foods did you give your baby, besides breastmilk?"
- **Order prompts in a sequence that is logical and builds trust**. Start with non-sensitive, descriptive questions, such as "Can you describe a typical day with your child?" before progressing to personal or emotionally charged topics — for example, "Have there been times when you worried about your child's health?"
- **Avoid leading or judgmental phrasing that could bias responses:**
  - In qualitative research, avoid questions framed as "Why didn't you…" Instead, consider "What are some things that might make it difficult to…?".
  - In quantitative research, use neutral Likert scale statements, as opposed to asking yes or no questions that imply a "right" answer (e.g., "Do you trust vaccines?"). For example, "I feel confident that vaccines recommended by health workers are safe," with responses organized by a scale of strongly agree, agree, neutral, disagree, and strongly disagree.
- **Incorporate probes, visual aids, or examples if needed**. For example, if inquiring about influences on decision-making, use a card sorting activity, or show images of common influencers (e.g., grandparents, religious leaders, health workers) and ask participants to explain who they listen to and why.

c. Facilitators and/or enumerators play a critical role in shaping how participants engage with the tools. Their ability to build trust, introduce activities clearly, and navigate sensitive topics with empathy can significantly influence the quality and depth of responses. Support these individuals through practical training to ensure consistency and reinforce ethical, respectful data collection.

d. Where possible, pilot tools informally with a few participants before a full rollout. This helps to surface confusing questions, adjust flow or language, and improve the overall clarity of research tools.

5. Ethics approval

Upholding ethics principles and standards is the responsibility of everyone involved in a research activity. For some activities an independent ethics review is required for assurance that the research design is appropriate and a thorough risk assessment has been completed.

Ethical review processes generally examine how participants are recruited; how data is collected, managed and stored; and how participants are supported after data collection concludes. For instance, ethics review boards will ensure that informed consent forms are written in clear, culturally appropriate language and conveyed in a way that ensures genuine understanding.

Drawing on standards elaborated in the **Belmont Report**[4] and Declaration of Helsinki, UNICEF's ethical principles guide decision-making across the project lifecycle to protect participants and communities. The key principles are:

- **Benefit**: Evidence generation must have a clear purpose that directly benefits communities while minimizing harm. Activities should be necessary and useful, prioritizing secondary data where appropriate.
- **Do No Harm**: A risk-informed approach must be employed to prevent harm, with a focus on children's best interests. Risk assessments, safeguarding plans, and data protection measures are essential to mitigating potential harm.
- **Respect**: Evidence activities should honor participants' dignity and well-being, ensuring informed consent, minimizing power imbalances, and protecting privacy. Localization and participatory approaches are encouraged, engaging stakeholders meaningfully and respectfully.
- **Inclusion**: All activities should be non-discriminatory, offering inclusive participation. Local expertise and contributions should be acknowledged, with careful consideration of biases and accessibility.
- **Fairness**: The burden and benefits of evidence activities should be balanced, with appropriate compensation for participants and sensitivity to cultural and community dynamics. Equity in outcomes and implementation is essential.
- **Integrity**: Adherence to UNICEF's core values and professional standards is critical. Conflict of interest must be managed, and project partners held accountable for maintaining ethical standards.
- **Accountability**: There must be clear accountability for decisions and actions throughout the project lifecycle. This includes transparent reporting of harms, continuous learning, and ethical publication of findings.

These principles emphasize a proactive, thoughtful approach to ethical compliance, ensuring that the rights and dignity of children and communities are protected throughout evidence activities."

4 U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research, The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research (Washington, DC: U.S. Government Printing Office, 1979), https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html (accessed June 19, 2025).

Behavioural research often requires engaging directly with individuals to understand their lived experiences, perceptions, and behaviours. Even when such engagement is not inherently sensitive, it may touch on deeply personal issues or involve vulnerable groups. For example, primary research on childhood vaccination may involve asking caregivers about trust in health services, past trauma, or difficult family dynamics—eliciting sensitive information even if unintentionally. In studies that involve sensitive topics, it may also be necessary to provide participants with a debrief or referral to relevant support services.

To determine whether formal ethics review (e.g., via an Institutional Review Board or national ethics committee) is required, teams should reflect on the following questions:

- Is the data being collected from or about people?
- Is the data identifiable (e.g., names, photographs, voices, or other traceable attributes)?
- Are any participants considered vulnerable (e.g., minors, displaced persons, survivors of violence, individuals with disabilities)?
- Does the research address sensitive topics (e.g., reproductive health, mental health, violence, illegal activity)?
- Could participation result in harm, distress, stigma, or legal or social risk?
- Will findings be shared externally (e.g., in donor reports, public presentations, or academic journals)?

If the answer to one or more of these questions is yes — or if there is uncertainty — teams are should consult a local ethics advisor or board. Even when formal approval is not mandated, a clearly defined internal ethics protocol should still be established to ensure informed consent, confidentiality, and participant protection.

See Table 13 on the following page for additional resources on ethical research.

TABLE 13. ETHICS RESOURCES

NAME	WHEN TO USE IT	SOURCE
UNICEF Procedure on Ethical Standards in Research, Evaluation, Data Collection and Analysis	To better understand procedures that establish minimum standards for UNICEF evaluations and ensure ethical oversight and accountability.	UNICEF
Ethical Considerations When Applying Behavioural Science in Projects Focused on Children	To better understand how the ethical behavioural science principles apply to projects that impact children. There is also a decision-support checklist for practitioners. The tool guides practitioners through key ethical decision points during an applied behavioural science project, and supports practitioners to identify when other perspectives may be needed.	UNICEF
Revised Common Rule Consent Forms	Guidance and templates for oral consent, remote consent (for studies conducted over the phone, for example), minors, etc. There is also a "basic elements of consent forms" checklist that is useful to review when creating consent forms for the primary research stage.	Johns Hopkins University
Basic course on Ethical Principles and Standards for Evidence Activities Involving People	This course presents ethics principles and standards for evidence activities, including practical examples.	UNICEF
IRB Research Plan for New Data Collection	Use it as a reference template when drafting a research protocol for any ethics or IRB submission, making sure to adapt it to your institution's forms and rules.	Johns Hopkins University
Research Plan Instructional Guide	Aids in explaining what an ethics review board will look for when reviewing a research plan.	Johns Hopkins University

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The Research Plan and Research Protocol were not developed by the original project team. They are recreated examples based on real project data and context.

Plan the research and develop a Research Protocol, including ethics IRB approval

Now that the project team was equipped with a strong understanding of the desired behavioural outcome — along with the key research objectives and questions — the project team formalized their research plan. Considering it was a sensitive topic that involved engaging with vulnerable populations (e.g., Syrian refugees), the project team decided to also develop a more comprehensive research protocol, based on the research plan.

For each of their research questions, the team determined the individuals required for conducting research, along with the appropriate research tools best suited for gathering information.

To support decision-making around participant sampling and research methods, the project team considered key constraints like time, budget, and potential behavioural biases that could have impacted the quality of their findings. They decided on a mixed-methods approach of qualitative and quantitative research, in addition to the recruitment of participants with a direct experience of vaccination services, and other key stakeholders.

D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Primary Research Plan CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

Pull the objectives and questions from previous activity and build on it to get to the primary research plan.

What to explore?

1. Capability: To identify caregivers' knowledge of gaps about the vaccine schedule and health worker communication.
2. Opportunity: To understand structural, logistical and social barriers to timely vaccination.
3. Motivation: To explore beliefs, emotions and trust factors influencing caregivers' intent to vaccinate on time.

What to ask?

1. What factors prevent caregivers from completing their children's vaccination schedules on time, even after a referral?
2. How do frontline health workers and outreach staff communication about follow-up appointments, and what barriers affect this communication?

Who?

1. Primary participants: 20-24 caregivers of under-/unvaccinated children
2. Secondary participants: Outreach workers (8-10), PHC staff (6-8), PHC managers and district health authorities (4-6), community leaders (4-6), MoPH stakeholders (4)

Sampling method: Purposive sampling to ensure representation; snowball sampling to reach undocumented or mobile refugee households. Variation captured: Gender, nationality, location, transportation access, outreach exposure

How?

1. In-depth interviews with caregivers, PHC staff and outreach workers to explore beliefs and motions.
2. Focus group discussions with community leaders and caregivers to uncover norms and collective experiences.
3. Structured observations of clinic visits and outreach interactions to observe real-world dynamics

When?

- Tools development and piloting: September 2017
- Fieldwork: September - December 2017
- Preliminary analysis: January 2018
- Stakeholder debrief and revision: February 2018
- Planning for prototyping: February - May 2018

Materials

Recording and note-taking equipment, including audio recorders. Accessibility and language aids, such as translated materials, visual cards for low-literacy participants

Roles and responsibilities

Project Lead (oversight, protocol review, partner coordination); Field Coordinator (recruit and manage facilitators); Research Assistant (transcribe, clean and code data); Arabic Translator (translate tools and transcripts)

Ethical considerations

IRB approval required
Protection of vulnerable populations: Special care for refugee households
Equity and inclusion: Prioritize hard-to-reach groups.

Overall, the Research Plan and Protocol proved to be critical throughout the project, helping the team to:

- Agree and align on their research plan
- Ensure there was clear documentation of their processes and actions, including any changes in course, objectives, or approaches
- Conduct their research with vulnerable populations and children in Lebanon, acting as a supporting document for ethics approval from the IR board

Research Plan

See sample research plan on the following pages.

Note: *While an actual Research Protocol would be longer and more comprehensive in its detail, for this case study we have developed some sample content of what could have been included in the Research Protocol. Please remember this is illustrative.*

RESEARCH PROTOCOL

Date: September, 2017
Study Contacts:
Principal Investigator: All Osseiran, Lead Behavioural Scientist, Nudge Lebanon
UNICEF Country Focal Point: Joanne Yammine, SBC Officer, UNICEF Lebanon
BIRD Lab Focal Point (for Study): Benjamin Hickler, Behavioural Scientist, UNICEF, BIRD Lab

1.0 Study Title:

Using behavioural insights to increase the demand for childhood vaccination in low resource settings

2.0 Background:

Lebanon has generally maintained high immunization coverage rates, nearing 90% for most routine vaccines (Diphtheria, Tetanus, Pertussis (DTaP), Hepatitis B, Oral Poliovirus Vaccine (OPV), and Haemophilus influenzae type B (HiB)). However, a rise in poverty levels, coupled with an influx of more than a million refugee in 2011 — many of whom were missing routine vaccines posed significant challenges to the national immunization program, which provides vaccination to more than half the children residing in Lebanon.

Evidence of weakening immunization coverage rates was first detected following a measles outbreak in 2013, and a drastic rise in the cases of mumps in 2015. This was validated by a district-based cluster survey conducted in 2016. The survey revealed pockets of low immunization coverage rates in several districts, among both refugee and host community households.

As a strengthening strategy to the national immunization program, the Ministry of Public Health (MoPH) Lebanon, in collaboration with United Nations Children's Fund (UNICEF), launched an accelerated immunization activities (AIA) programme in November 2017 to provide free quality immunization services for every child through the primary healthcare system. In addition to the support offered to primary healthcare centers (PHCs), the programme employed a community-based outreach approach to (a) identify un- or under-vaccinated children, (b) educate and raise awareness of caregivers on the importance of childhood vaccination, (c) collect children's vaccination records, (d) refer children missing vaccines to the nearest participating healthcare center, and (e) follow-up with caregivers to ensure that full immunization is achieved. Raising awareness, strengthening public healthcare system, and providing access to free quality services are necessary prerequisites to facilitate and enable the uptake of childhood vaccination. However, these measures are seldom enough to ensure sufficient demand from beneficiaries.

The target sample consisted of households with children aged 0-16 years who were falling behind their routine vaccination schedule according to Lebanon's national immunization calendar. The sample included households that were previously referred to a health care centre but failed

to maintain their children's vaccination up to date (follow-up visits), as well as newly identified households with un- or under-vaccinated children (outreach visits). Households with children who were up to date with their vaccination were excluded from the trial.

3.0 Study Objective(s) and Expected Outcomes:

Despite caregivers' best intentions, evidence from years of research in psychology and the behavioural sciences on vaccine uptake point to the presence of decision biases and social influences (henceforth behavioural barriers) that impact the motivation to vaccinate. The expected outcome of this research is to identify these behavioural barriers and address them using the appropriate behaviour change techniques is critical for the success of any campaign or activity seeking to boost national immunization rates.

Following the COM-B model, the study objectives are:

Capability

- To identify caregivers' knowledge and gaps in caregivers' understanding of the vaccination schedule, including timing, number of doses, and next steps after AIA outreach visits.
- To explore whether PHC staff and outreach workers have the knowledge and confidence to consistently communicate accurate, understandable vaccination information.

Opportunity

- To understand the physical, logistical, and structural barriers (e.g., clinic accessibility, mobility constraints, service hours) that affect caregivers' ability to vaccinate.
- To examine how social norms, household dynamics, and community beliefs shape caregivers' decisions to vaccinate.

Motivation

- To explore the beliefs, emotions, and values that influence caregivers' perceived urgency or importance of completing all recommended vaccinations on time.
- To identify factors contributing to trust or mistrust in vaccination services, and how these perceptions impact motivation to act on outreach or PHC guidance.

3.1 Research Questions:

Primary Research Questions:

1. What factors prevent caregivers from completing their children's vaccination schedules on time, even after a referral?
2. How do frontline health workers and outreach staff communicate about follow-up appointments, and what barriers affect this communication?

4.0 Study Design/Methodology.

Given the complexity of vaccination behaviour, a mix of qualitative and observational method include:

Method	Purpose
In-depth interviews	To explore individual beliefs, emotional drivers, and decision-making processes related to vaccination. These will capture caregivers' personal experiences, PHC staff's communication practices, and outreach workers' perspectives.
Focus group discussions	To examine collective perceptions, social norms, group dynamics, and shared community experiences. FGDs with caregivers, leaders, and influencers will provide insights into the wider social and cultural context influencing vaccine uptake.
Structured field observations	To capture real-time interactions at PHCs and during household visits, documenting behaviours, communication styles, and systemic bottlenecks in service delivery. Observations will also provide contextual data on household dynamics and community settings.

Rationale of these methods

- Interviews will capture nuanced caregiver motivations and emotional contexts, especially for refugees who may distrust formal systems.
- FGDs will enable us to gather insights at scale and support group-based activities (e.g. community network mapping with participants). However, this method will not be used for more sensitive topics (e.g. gender norms within the refugee communities), or potential issues like discrimination at health clinics.
- Observations at vaccine clinics will document the vaccination process, including system bottlenecks or behavioural patterns not captured in verbal responses. Observations within the community will document information on the interpersonal relationships between outreach workers and caregivers, as well as community and household dynamics and living conditions.
- Where feasible, participant journaling and visual aids (vignettes or calendars) may supplement verbal recall.

Integration of Methods

The use of multiple qualitative and observational methods will enable triangulation, strengthening the reliability and validity of findings. For example, themes emerging in interviews will be compared with group dynamics observed in FGDs and behaviours documented through field observations. This layered approach ensures that both individual and systemic perspectives are represented.

4.1 Sampling Criteria and Recruitment:

Sampling Technique

A purposive sampling strategy will be applied to ensure that participants reflect the characteristics most relevant to the study objectives, particularly households with un- or under-vaccinated children. For harder-to-reach populations, especially undocumented or highly mobile refugee caregivers, snowball sampling will complement recruitment by drawing on trusted community actors and outreach teams. This mixed approach will maximise diversity while ensuring access to populations at heightened risk of incomplete vaccination.

Inclusion Criteria

- Caregivers (mothers, fathers, grandparents, or other guardians) of children aged 0-6 years who are un- or under-vaccinated according to Lebanon's national immunization calendar.
- Households that were previously referred to a PHC but failed to complete vaccination schedules, as well as newly identified households with missed doses.
- Outreach workers, PHC staff, PHC managers, community leaders/influencers, and Ministry of Public Health (MoPH) representatives directly involved in vaccination demand and delivery.

Exclusion Criteria

- Households where children are fully up to date with vaccinations.
- Caregivers who are unwilling or unable to provide informed consent
- Individuals under the age of 18, unless participating as caregivers of younger siblings and with guardian consent.

Sample Sizes and Rationale

The sample sizes are designed to balance depth and breadth: large enough to capture variation across refugee and host populations, while remaining feasible within the project timeframe.

Primary participants (caregivers):

20-24 caregivers across seven districts. This sample is sufficient to reach thematic saturation in qualitative interviews, while allowing variation by gender, nationality, geography, and household

mobility.

Secondary participants:

- 8-10 outreach workers (to explore household engagement practices).
- 6-8 PHC nurses and staff (to capture communication strategies at service delivery points).
- 4-6 PHC managers/district-level health authorities (to understand system-level challenges).
- 4-6 community leaders or influencers (to capture community and normative influences).
- 4 MoPH stakeholders (to reflect policy-level perspectives).

Variation to Capture

Recruitment will aim to ensure diversity across:

- Gender (mothers and fathers as decision-makers).
- Nationality (Syrian refugees, Lebanese host communities, and other minority groups where present).
- Geography (urban vs. peri-urban districts).
- Mobility (households with and without access to transport).
- Exposure (households previously visited by outreach workers vs. those not reached). This variation will allow the study to map behavioural barriers and enablers across different household types and community contexts.

Recruitment Procedures

Participants will be identified in collaboration with PHCs, outreach workers, and community-based organisations. Outreach staff will provide referrals to eligible households, and snowball techniques will be used where initial access is limited. Recruitment will emphasise voluntary participation, with clear communication that refusal will not affect access to services.

4.2 Consent/Assent

Informed consent: All participants will be provided with an overview of the study, including purpose, confidentiality, and their right to withdraw. Participants will be asked to sign a consent form or provide recorded verbal consent.

4.3 Methods of Data Collection:

This study will employ a combination of qualitative and observational methods, each carefully matched to the research questions and designed using behavioural insights best practices. The

primary methods include semi-structured interviews, focus group discussions (FGDs), and structured observations in both community and healthcare facility settings.

Data Collection Methods and Tools

- **Semi-structured interviews:** One-on-one interviews with caregivers, PHC staff, and outreach workers will be conducted using COM-B aligned guides. These guides will probe knowledge, opportunity, and motivational factors, while allowing flexibility to capture nuanced beliefs, emotions, and decision-making processes.
- **Visual prompts and vignettes:** To support recall and depersonalise sensitive issues (e.g., mistrust, shame, fear), visual materials and hypothetical vignettes will be incorporated into interview and FGD tools. These approaches reduce social desirability bias and make participation more engaging, especially for low-literacy participants.
- **Focus group discussions (FGDs):** FGDs will be conducted with community leaders, caregivers, and influencers to capture shared experiences, social norms, and group dynamics. Facilitation guides will ensure consistency across groups, with built-in exercises (e.g., community mapping) to encourage interactive dialogue.
- **Structured observations:** Field researchers will use observation checklists to systematically document PHC interactions (e.g., communication at registration and vaccination points) and outreach household visits (e.g., referral processes, family responses, social dynamics). Photos and short video recordings may be taken to capture contextual details, with participants' consent, and all images will be anonymised before analysis.

Data Collectors

Research activities will be implemented by trained facilitators sub-contracted by Nudge Lebanon and UNICEF Lebanon. These facilitators will include both male and female researchers to enable culturally sensitive engagement with diverse participants. All field researchers will sign confidentiality agreements and adhere to safeguarding protocols.

Training and Capacity Building

To ensure correct and consistent application of the methodology, all data collectors will undergo a structured training process. Three in-person workshops will be held at least one week prior to fieldwork. Training will combine:

- Theoretical components (study objectives, behavioural science framing, safeguarding, and informed consent procedures).
- Practical exercises (role-play of interviews and FGDs, mock observations, use of visual prompts and checklists).
- Safeguarding orientation (how to handle disclosures of distress, sensitive issues, or safeguarding concerns in line with UNICEF's protocols).

Supervision and Quality Assurance

- Daily debriefings: At the end of each research day, virtual debrief sessions will be held with field researchers and the project team to review emerging findings, identify challenges, and reinforce adherence to protocols.
- Spot checks: Senior researchers from Nudge Lebanon will accompany a subset of interviews, FGDs, and observations to ensure fidelity to the guides and ethical standards.
- Documentation: Research guides and annotation sheets will be provided to all facilitators to capture observations consistently and to supplement audio recordings.

4.4 Data Entry and Analysis:

A thematic analysis will be conducted using the COM-B framework. Coding will be structured around capability (e.g., awareness of schedule, recall accuracy), opportunity (e.g., access, system navigation, social expectations), and motivation (e.g., fear, shame, priorities, trust in the health system). An additional tagging of behavioural profiles will identify participants with similar behaviours, demographics, beliefs, and needs. Triangulation will be used to validate self-reported behaviour with observations and facilitator notes. Comparative analyses will be conducted across refugee versus host populations and across different PHCs/districts. NVivo or Dedoose will be used for qualitative coding, and social network data will be mapped using matrix visualisations.

Data Handling Procedures

- **Transcription**: Interviews and focus groups will be transcribed in Arabic and/or French and then translated to English by trained research assistants contracted for this study. Transcribers will sign confidentiality agreements.
- **Coding**: Coding of transcripts and field notes will be conducted by members of the research team trained in qualitative methods, using NVivo or Dedoose. Double-coding will be applied to a sample of transcripts to ensure reliability.
- **De-identification:** All transcripts will be de-identified by removing names and other direct identifiers during the transcription process. Participants will be assigned unique study IDs. The key linking IDs to participant information will be stored separately on a secure server accessible only to the project team members.
- **Storage/Transfer**: Electronic data will be stored on UNICEF's secure, password-protected servers. Files transferred between team members will use encrypted channels (SharePoint). Physical field notes, if used, will be scanned, uploaded, and then stored in a locked cabinet at the local partner office.
- **Access**: Access to data will be restricted to authorised members of the study team. Each team member will use individual login credentials and the unique password to access the documents. Access rights will be granted and monitored by the principal investigator.

- **Archiving**: At the end of the study, de-identified datasets and final codebooks will be archived on UNICEF's secure servers for the 5 years. Any personal identifiers will be permanently deleted after analysis is complete.

5.0 Data Protection and Study Participant Confidentiality

- **Data privacy:** No personally identifiable information will be recorded. The faces of people in photos and videos will be blurred.
- **Sensitive topics:** Interviews may explore emotional topics like shame, confusion, or mistrust. Trauma-informed approaches and vignettes will be used where appropriate. As many participants are considered vulnerable populations as refugees, additional care will be taken, including ethics training for researchers and providing participants with links to further follow-up psychosocial support, if needed.
- **IRB Review:** The protocol will be submitted to an institutional review board (IRB) for ethical clearance.
- **Equity:** Care will be taken to include hard to reach and underrepresented voices, particularly female-headed households or households where caregivers are not literate or digitally connected.
- **Compliance with local requirements:** All study procedures will comply with applicable data protection laws and regulations in Lebanon. Access to participant data will be limited to authorised study personnel only, and all reasonable safeguards will be taken to ensure confidentiality, in accordance with local ethical and legal requirements.

5.1 Risk Assessment and Management:

Participation in this study is considered to pose minimal risk. However, given the study's focus on vulnerable populations such as refugees and low-income households, careful risk management measures will be put in place.

Potential Risks and Harms

- **Emotional distress:** Discussions on vaccination decisions may surface feelings of guilt, shame, fear, or mistrust of the health system.
- **Social risks:** Caregivers may fear stigma for admitting lapses in vaccination or expressing mistrust of healthcare providers.
- **Privacy concerns:** As refugees may be undocumented or marginalised, participants may worry about misuse of information or unintended disclosure.
- **Power dynamics:** Caregivers may feel pressure to participate if approached by outreach workers or community leaders.

7.0 Findings Dissemination

Findings from this study will be disseminated in a way that ensures transparency, respect for participants, and alignment with UNICEF and Ministry of Public Health (MoPH) standards.

Dissemination to Stakeholders

- Preliminary results will be presented to the MoPH, UNICEF Lebanon, and Nudge Lebanon in a joint debrief session to validate findings and refine interpretation.
- A final report will be shared with national stakeholders, including PHC managers and outreach supervisors, to inform programming decisions and potential prototype development.
- A policy brief will be prepared for decision-makers highlighting key behavioural barriers and recommendations.

Dissemination to Participants and Communities

- A simplified summary of findings will be developed in Arabic, using clear and accessible language. This will be shared through community meetings facilitated by outreach workers and community leaders.
- Where appropriate, visual posters or leaflets with key insights and recommendations (e.g., service accessibility improvements, communication messages) will be made available at participating PHCs.
- No individual or household data will be shared; only de-identified, aggregate findings will be communicated back.

Publication Rights

- The investigators retain the right to publish study results in academic or practitioner journals, subject to UNICEF and MoPH review and approval prior to submission.
- Any publication will ensure that individual participants cannot be identified, and only de-identified data will be used.
- Authorship and acknowledgement will follow internationally accepted standards, recognising the contributions of UNICEF, MoPH Lebanon, Nudge Lebanon, and the research team.

Risk Management Plan

- **Informed consent:** Participants will be provided clear, simple-language information about the voluntary nature of the study, confidentiality measures, and their right to withdraw at any time without consequence.
- **Safeguarding training:** All facilitators and research staff will undergo training on safeguarding, trauma-informed interviewing, and UNICEF's child protection protocols prior to data collection. Referral pathways: If a participant shows signs of distress, or discloses sensitive safeguarding issues (e.g., neglect, abuse, exploitation), field staff will follow UNICEF Lebanon safeguarding protocols and provide referral information to local psychosocial support or protection services.
- **Minimising distress:** Sensitive questions will be asked using vignettes or indirect framing to depersonalise experiences where appropriate. Researchers will be trained to stop or pause an interview if a participant shows discomfort. • Privacy and confidentiality: All data will be de-identified, and no personally identifiable information will be disclosed. Photos or videos taken for observation purposes will have faces blurred.
- **Equity safeguards:** Special attention will be given to ensuring that marginalised groups (e.g., female-headed households, illiterate caregivers) are included and treated respectfully. No group will be disadvantaged by participation.

Oversight and Monitoring

The Principal Investigator will be responsible for monitoring risks and safeguarding adherence throughout the study. Daily debriefings with field teams will include a review of any safeguarding issues or participant concerns. Any incidents will be documented, reported to UNICEF's safeguarding focal point, and addressed promptly.

6.0 Timetable

Outline key timeline for this study (add/delete rows as needed)

ACTIVITY (e.g., pilot test tools)	WHEN (month/year)
Tool development and piloting	Sept.–Oct 2017
Training of field researchers and safeguard orientation	October 2017
Exploratory primary data collection (fieldwork in PHCs and communities)	Nov.–Dec. 2017
Transcription, translation, and preliminary analysis	Dec. 2017 – Jan. 2018
Stakeholder debrief and joint validation workshop (MoPH, UNICEF, Nudge Lebanon)	January 2018
Revision of findings	February 2018
Drafting and submission of final study report to UNICEF.MoPH	March 2018

STEP 4:

Conduct the primary research

In this step:

This step entails preparing for and carrying out primary data collection in the field, including finalising logistics, training the team, recruiting participants, and implementing the research plan with rigour and flexibility.

Why it matters:

Strong fieldwork is essential for producing behavioural insights that are rich, reliable, and relevant. Without proper preparation and oversight, critical voices may be missed, data quality may suffer, and ethical risks may increase. This step ensures that the research is thoughtfully planned and sensitively conducted, to ensure inclusion, accuracy, and ethical integrity.

How to do it:

1. Prepare for field work

Before heading into the field, it's important to ensure that the research is well-planned and set up for success. Confirm the following steps and procedures are in place:

a. **Field team have clear roles and responsibilities:** Each team member should have a specific role, such as recruiter, enumerator, facilitator, translator, or note taker. When possible, matching interviewers to participants based on gender, language, or cultural background can increase comfort and trust. This is especially important when working with marginalized groups or sensitive topics. For example, mothers may prefer to be interviewed by fellow mothers.

b. **The field team is trained on tools and ethical guidelines:** All team members should be trained on the research tools, including how to use the guides, ask follow-up questions, and record responses accurately. Make sure enumerators or facilitators are trained to probe beyond the broad, high-level action the project is trying to influence (e.g., getting vaccinated), in order to uncover the smaller, often overlooked steps and decisions that shape whether a behaviour happens. For example, researchers should consider whether and how caretakers plan ahead, discuss with family, and remember appointments— smaller decisions or actions that bridge the gap between intention and follow-through. Brief the field team to listen for unexpected behavioural insights (e.g., unanticipated motivators, unusual influencers, or surprising points of friction) and record them, even if they fall outside the discussion guide. Ethical training should cover informed consent, confidentiality, respectful engagement, and handling sensitive topics and disclosures.

c. **Research participants are recruited according to the sampling strategy** (see Table 10. Sampling strategies for more detail): This may include building simple screening tools, such as a brief checklist or set of questions to confirm key criteria like age, caregiver status, geographic location, or vaccination history, to ensure participants meet the required characteristics. For example, if the study focuses on caregivers of children under five who missed routine vaccinations, the screener might ask: "Do you have a child under the age of five?" and "Has your child missed any routine immunization appointments in the past year?" Careful recruitment is especially important when aiming to include underserved or underrepresented groups.

d. **Data collection is scheduled appropriately:** Fieldwork should be scheduled with sensitivity to local routines, customs, and seasonal events. Special care should be taken to avoid clashing with work, caregiving duties, or religious observances, especially for women or other groups with limited free time.

e. **Appropriate locations are selected:** The interview or discussion setting should feel safe and comfortable for participants.Choosing the right location helps to create an environment where participants can speak openly and without fear of judgment or interruption. For example, interviews with adolescent girls may be better conducted

in schools or health centres, rather than at home where family members are present.

f. **Necessary approvals are granted:** Fieldwork should only proceed once the required permissions and approvals are in place. This may include IRB or ethics committee approval, as well as formal or informal authorization from local authorities and community leaders.

g. **Research tools are tested:** Research tools should be tested before fieldwork begins to ensure they are clear, culturally appropriate, and aligned with the research objectives.

h. **Other logistical arrangements are in place:** Teams should also confirm that materials, devices, transport, and any translation support are in place.

i. **Security and safety protocols are developed:** Before fieldwork begins, researchers should ensure all team members are aware of relevant safety procedures, especially when working in hard to reach or conflict-affected areas. This includes protocols for communication, transportation, emergency contacts, and actions to take if conditions change unexpectedly. Briefings should also cover cultural norms, local sensitivities, and any restrictions that could affect team mobility or participant comfort.

2. Collect data

Once fieldwork begins, it's important to monitor the quality of work and respect for participants. While the field team leads day-to-day interactions, the project team should provide support by promoting good practices and addressing challenges as they arise. Key best practices to reinforce with the team include:

a. **Informed consent is given by participants:** Participants must understand the purpose of the research, how their information will be used, and what their rights are. Consent can be verbal or written, depending on context. If recording audio, taking photographs, or capturing video, additional consent must be obtained specifically for those activities.

b. **Space is created for trust and openness:** The best insights often come when participants feel safe and unjudged.Team should be encouraged to take time at the start of each interaction to build rapport, explain the purpose of the work clearly, and listen with curiosity, rather than rushing through questions.

c. **Power dynamics are respected:** Field teams should be trained to recognize who holds power in group settings and adapt facilitation accordingly. In qualitative interviews, attention should be paid to gender, age, and social dynamics that may influence who speaks and how freely.

d. **Performative or expected responses are considered:** Particularly in group settings or when interviewing community members used to programme visits, participants may give answers they think are expected. Teams should be encouraged to probe gently but specifically to unpack underlying motivations or contradictions.

e. **Documenting what's not said:** Note takers and facilitators should be encouraged to go beyond quotes. Capturing tone, hesitation, body language, or contradictions can add critical context when synthesising social and behavioural data. Non-verbal cues are especially important when working across language or cultural differences.

f. **Daily debriefs are built in:** Teams should come together at the end of each day to discuss what's working, what feels off, and what unexpected themes are emerging. Debriefs are a key moment to adjust tools, refine probes, and align on what to explore further the next day.

STEP 5:

Analyse and synthesise findings

In this step:

The goal of this step is to shift from raw findings to a clear, coherent understanding of what enables, drives, or inhibits the behaviour in question.

Three different behavioural frameworks frame the analysis of the data to identify insights:

- **Behavioural Profiles** illuminate who the key actors are, what shapes their choices, and what challenges they navigate.
- **Behavioural Mapping** and **Diagnoses** trace the micro-behaviours along the pathway to behaviour, and identify barriers and enablers behind these micro-behaviours.

These tools provide the foundations for designing interventions, and will guide the solutions created in the next phase.

Associated Tools:

- Behavioural Profile (optional)
- Behavioural Map and Diagnosis
- Feasibility–Impact Matrix

Why it matters:

Insight doesn't come from data alone. Rather, how we make sense of this data brings findings into a coherent behavioural narrative and clarifies not only what is occurring, but why.

Critically, this step anchors the work at hand in the lived experiences of real people. **Behavioural profiles** bring actors into focus. **Behavioural maps** trace the small, often unnoticed actions that make or break outcomes. And **behavioural diagnoses** expose where behaviour breaks down and what stands in the way.

Done well, this step will accomplish the following:

- **Identify strategic entry points for change.** For example, data could reveal how vaccination drop-off rates increase after the first shot, with no follow-up system currently in place within the health system.
- **Distill complexity into clear, actionable patterns.** For example, both quantitative and qualitative data could reveal how caregivers tend to miss vaccination appointments on days where there is extreme weather.
- **Ensure designs respond to real-world constraints—not assumptions.** For example, after noting how some refugee mothers rely on male relatives who are only available on weekends for transport to the clinic, consider how weekday appointments may not be feasible for some households.
- **Build the behavioural foundation for solutions to be prototyped next.** For instance, after consulting findings from both caregivers and health care workers who noted a lack of communication from the health system, a potential solution could be designing a follow-up message that reminds caregivers of the date of the next scheduled vaccine.

This ensures a focused, evidence-based understanding of where change is needed, and what is required to make it possible, when designing interventions.

How to do it:

1. Clean the data

After primary research, data is rarely ready for immediate use. For example, transcripts may be incomplete, notes may need clarification, and inconsistencies may need to be resolved before meaningful patterns can emerge. Before any analysis can begin, raw data needs to be cleaned, coded, and sense-checked. Two helpful guides for navigating this stage include:

- The SAGE handbook of Qualitative Data Analysis
- The World Bank's DIME Handbook (focused on rigorous, large-scale data collection)

When transcribing interviews or focus group discussions, AI tools can speed up the process significantly, though human review should be used to ensure accuracy

and nuance, especially for non-standard dialects or sensitive topics.Well-reviewed transcription software includes Auris, Trint, TurboScribe, and Descript.

2. Conduct quantitative analysis

Quantitative data,such as survey responses, service records, or administrative datasets, should be analysed in a structured way to identify patterns and trends at scale. This type of analysis can help answer key questions such as:

- How common is the target behaviour among the population of interest?
- How do behaviours vary across different groups (e.g. by gender, location, caregiver type)?
- Are specific attitudes or access barriers associated with greater or lesser likelihood of completing the behaviour?
- What proportion of individuals drop off at key points in the service journey?

Quantitative analysis might include frequency counts, cross-tabulations, correlations, or regression analysis, depending on the dataset and available skills. Where relevant, survey questions should be mapped to COM-B domains to interpret findings through a behavioural lens.

For existing datasets (e.g. vaccination records or attendance logs), descriptive statistics are often enough to highlight trends and flag where behavioural bottlenecks occur.

See the Learn more section for guidance on basic statistical methods, data cleaning, and survey-based research.

3. Organize and tag qualitative data

Qualitative data (e.g., interviews, focus groups, observations) should be organized and tagged it in a way that supports structured analysis. Tagging means linking observations or quotes from participants to the behavioural drivers that might be influencing their actions, using a framework like COM-B.

a. Start by sorting transcripts, notes, or excerpts based on who the data is regarding (e.g., caregivers, providers, influencers) and what behaviour it relates to (e.g., completing vaccination, seeking ANC).

b. Create a simple coding matrix to help structure the analysis, for example by mapping behavioural drivers or barriers against COM-B domains. Table 14 provides an excerpt of coding matrix using a routine immunization example:

TABLE 14. CODING MATRIX

COM-B SUB-DOMAIN	DRIVER OR BARRIER	TAG
Physical capability	Barrier	Limited physical mobility
Psychological capability	Barrier	Competing priorities
	Barrier	Limited knowledge of vaccine schedules
	Driver	Knowledge of vaccine benefits
Physical opportunity	Barrier	Distance to the clinic
	Barrier	Unfavourable clinic hours
	Driver	Bundled services
Social opportunity	Barrier	Unfavourable descriptive norms
	Driver	Influence of mother-in-law
Reflective motivation	Barrier	Fear of side effects
	Driver	Belief in vaccine efficacy
Automatic motivation	Barrier	Present bias
	Barrier	Emotional overload
	Driver	Helpful habit or routine

Sort through the data (transcripts, notes, quotes) and tag excerpts using specific tags, not only the broad domain. A high level of detail helps to surface patterns, contradictions, and priority barriers that are directly useful for design. For example:

- Quote: "I didn't go back because I thought one shot was enough." → Tag: Limited knowledge of vaccine schedules
- Quote: "My neighbour told me her baby got sick after the vaccine." → Tag: Fear of side effects
- Quote: "I ask my mother-in-law for advice regarding my child' health." → Tag: Influence of mother-in-law

Use a spreadsheet, shared document, or qualitative coding software such as Dedoose, NVivo, or Atlas.ti to organise and tag data. If working manually, colour-coding or margin labels can support quick tagging. If using a cloud-based version of these tools, ensure that any data sharing complies with UNICEF data protection policies and informed consent agreements.

4. Create behavioural profiles (optional)

This step is particularly useful if the research — both desk research and primary data — generates rich detail about the people behind the explored behaviours. This means obtaining insight not only into what people do, but why, in what context, and with what challenges.

A behavioural profile is a tool used to synthesise insights about the capability, opportunity, and motivation-based drivers behind a specific behaviour or set of behaviours within a target population. While other similar tools such as user personas or demographic segmentation focus on demographic traits or broader needs, a behavioural profile is specifically designed to uncover the "why" behind behaviours, drawing on frameworks like COM-B.

Developing a behavioural profile is especially helpful when an intervention requires a deep understanding of behavioural drivers in order to be effective. It can also be useful if there are people who can be differentiated by a set of specific behavioural barriers, habits, emotions, and socioecological context. For example, two caregivers might look similar demographically (they are both 30-year-old mothers living in rural areas with two young children) but their behavioural profiles differ: one delays vaccination because she believes her child is not at risk and fears side effects (low motivation), while the other fully intends to vaccinate but misses appointments due to long travel distances and lack of childcare (low opportunity).

While a user persona might group these women together based on age, location, and caregiving role, behavioural profiles would separate them based on the different drivers influencing their behaviours.

For each behavioural profile, aim for at least 15 participants that share a common set of traits and behavioural features. It is important to highlight that 15 participants as a total sample size of the research might be insufficient if there is high heterogeneity — i.e., high variability in participants' profiles and responses.

TABLE 15. TYPE OF INFORMATION REQUIRED TO DEVELOP BEHAVIOURAL PROFILES AND POTENTIAL SOURCES OF INFORMATION

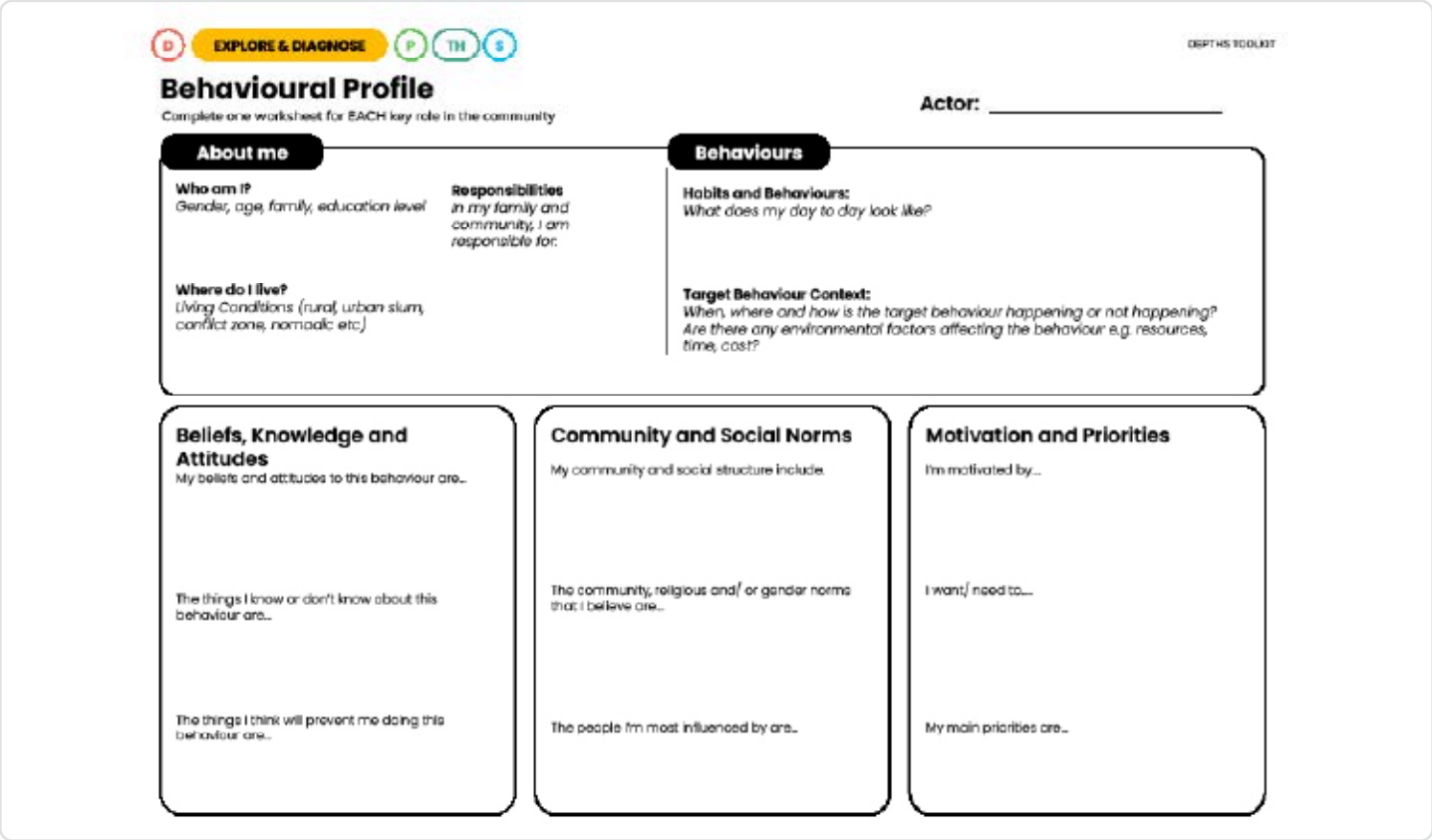
TYPE OF INFORMATION	POTENTIAL SOURCES
Basic Demographic Information: Data such as age, gender, education, household type, income level, geographic location, role in their community (e.g., caregiver, healthcare worker, religious leader, etc.). These data provide necessary context and help to distinguish how different populations experience behavioural drivers differently.	Stakeholder and audience map (Step 1 — Define phase): Review of existing administrative data that was used to build the stakeholder and audience map. Initial desk research (Step 1 — Explore & Diagnose phase): Administrative data such as data sets from the Ministry of Health, existing data sets from UNICEF, etc. Primary research (Step 4 — Explore & Diagnose phase): Data collected through surveys.
In-depth behavioural information that can be structured around a behavioural model (e.g., COM-B mode): Coded, grouped, or analysed using a framework like COM-B, this includes data on people's capability (knowledge, skills), opportunity (social and physical context), and motivation (beliefs, emotions, intentions) related to the target behaviour.	**Plan primary research (Step 3 — *Explore & Diagnose* phase):** A pre-defined behavioural model to analyse the data collected during the primary research. For example, the COM-B model. **Primary research (Step 4 — *Explore & Diagnose* phase):** Qualitative information typically gathered through in-depth interviews, focus group discussions, and/or observations, which help to identify barriers and enablers related to the target behaviour.

Unlike fictional personas, these profiles are rooted in evidence. They help to move beyond abstract categories like “caregiver” or “youth” by giving form to real people navigating real decisions. Most importantly, they help to shape intervention strategies by illuminating actor-specific needs and the contexts in which decisions are made. Here are steps to take to create behavioural profiles:

a. Review data from the fieldwork. While scanning the information, identify groups of actors who share similar behaviours, beliefs, influences, or experiences. These groups may be defined by role (e.g., frontline nurse, adolescent girl), life stage (e.g., first-time mother), or social position (e.g., respected elder, decision-maker in the home). If there are meaningful differences within a group, such as confidence levels, decision-making power, or social standing, this may require creating more than one profile.

b. Synthesise key findings about the background information, typical behaviours, motivations, knowledge, and norms relating to each profile. Use the Behavioural Profile worksheet to structure these findings.When planning the development of behavioural profiles, consider adding additional research questions to the research plan to learn more about the key actors being researched. For each profile, synthesise the following in the worksheet:

- **About me:** Key demographic information about who the person is and their living conditions.
- **Behaviours:** Information on the typical day-to-day behaviours and habits of the person, as well as any contextual information on when, where, and how they would conduct the target behaviour.
- **Beliefs, Knowledge and Attitudes:** What the person believes, knows and thinks about the target behaviour.
- **Community and Social Norms:** Overview of the community and social structures of this person, the religious, cultural, and gender norms influencing them, and who influences them.



D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Behavioural Profile

Complete one worksheet for EACH key role in the community

Actor: _____

About me

Who am I?
Gender, age, family, education level

Responsibilities
In my family and community, I am responsible for:

Where do I live?
Living Conditions (rural, urban slum, conflict zone, nomadic etc.)

Behaviours

Habits and Behaviours:
What does my day to day look like?

Target Behaviour Context:
When, where and how is the target behaviour happening or not happening? Are there any environmental factors affecting the behaviour e.g. resources, time, cost?

Beliefs, Knowledge and Attitudes
My beliefs and attitudes to this behaviour are...

The things I know or don't know about this behaviour are...

The things I think will prevent me doing this behaviour are...

Community and Social Norms
My community and social structure include

The community, religious and/ or gender norms that I believe are...

The people I'm most influenced by are...

Motivation and Priorities
I'm motivated by...

I want/ need to...

My main priorities are...

- **Motivation and Priorities:** The key motivations, priorities, wants, and needs driving this person's decision-making and behaviours.

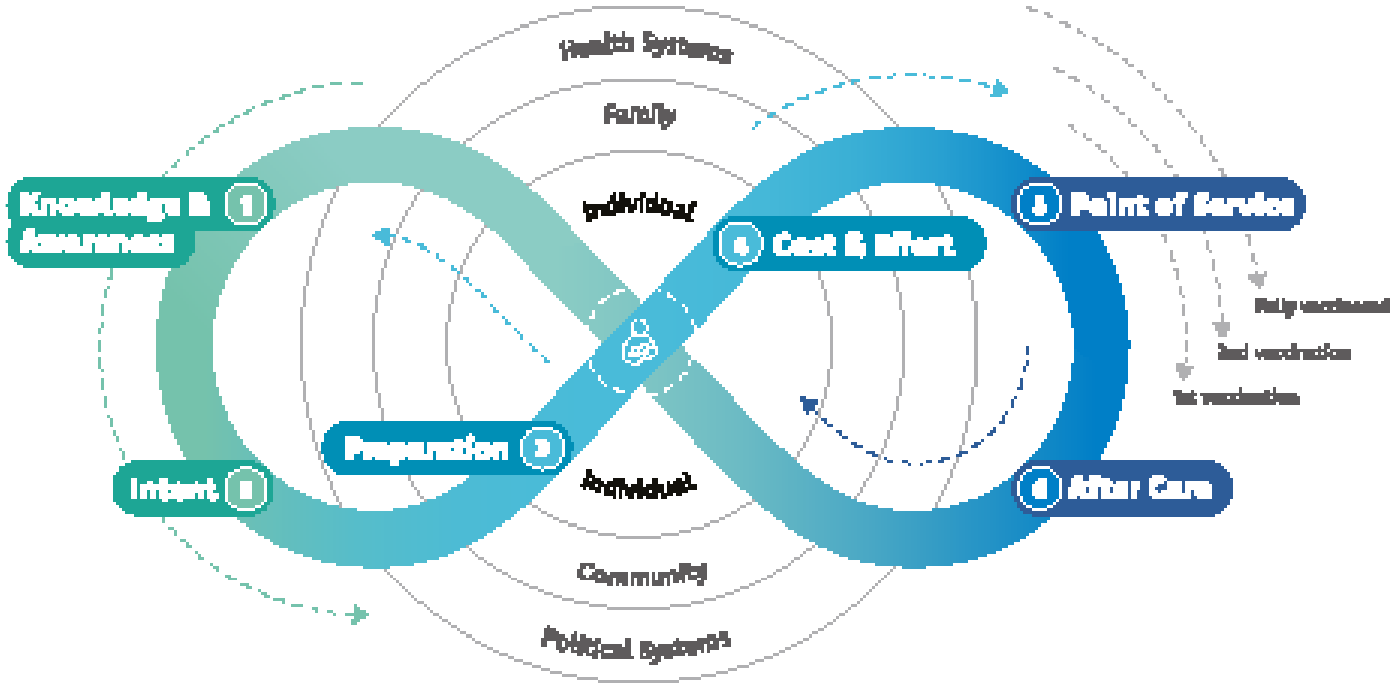
When building behavioural profiles, keep the following note in mind that **a strong behavioural profile should evoke empathy.** It should feel like a real person — someone whose daily realities, values, and constraints are understood. When done well, these profiles become more than summaries; they serve as living reference points throughout the process. Return to them during behavioural mapping, diagnosis, and design to stay grounded in the lives and contexts of the people the work is meant to serve.

5. Build a behavioural map and diagnosis

The Behavioural Map and Diagnosis worksheet helps to break down complex behaviours into smaller, observable steps: micro-behaviours. Rather than looking at a behaviour in isolation (e.g. "getting a child vaccinated"), this worksheet maps out the full behavioural journey: what happens before, during, and after the key action. For each moment in the journey, the tool helps to compare the ideal behaviour (what will ideally occur) with the real behaviour (what actually occurs). The tool then helps to identify the barriers and enablers that hinder or support the desired behaviour, along the different steps of this journey. By visualising the journey in this way, teams can spot the specific points where people get stuck, delayed, or drop off, and begin to design solutions targeted at those points.

Similar tools, such as user journeys and service journeys, are used in human-centred design and service design. These tools can vary in complexity: from quick, informal sketches to detailed, multi-actor maps. Some have even evolved into broader frameworks used across sectors to diagnose behavioural challenges.

One example is the Journey to Health and Immunization developed by UNICEF, which outlines the sequence of



decisions and actions a caregiver must navigate to access and complete routine immunization or other health services. Common Thread also offers a simple template and guide on how to develop behavioural maps, as well as the Center for Advanced Hindsight from Duke University.

To develop a behavioural map and diagnosis:

a. Identify the key profile (e.g. caregiver, health worker, adolescent) whose behaviour is the focus, and write it down. If behavioural profiles have already been developed for this actor, use it to map and tailor the journey specifically to that profile.

b. Define the target behaviour, or the specific action of interest (e.g. completing the full immunization schedule, registering a birth, practicing exclusive breastfeeding in the first 6 months, enrolling and regularly attending early childhood education, consistently using a latrine rather than open defecation).

c. Next, map the journey moments. These will be the smaller, observable steps that lead to or surround the target behaviour. These steps should capture what happens before, during, and after the target behaviour.

Table 16 on the following page provides illustrative examples of journey moments, across a range of behaviours linked to child rights outcomes. While not exhaustive, the list is intended to guide teams in thinking through the kinds of micro-behaviours that may be relevant to their context. It can be adapted based on the specific actor, behaviour, and setting.

TABLE 16. EXAMPLE JOURNEY MOMENTS ACROSS BEHAVIOURAL AREAS

COMPLETING THE FULL IMMUNIZATION SCHEDULE	REGISTERING A BIRTH IN THE FIRST 6 MONTHS	PRACTICING EXCLUSIVE BREASTFEEDING IN THE FIRST 6 MONTHS	ENROLLING AND REGULARLY ATTENDING EARLY CHILDHOOD EDUCATION	CONSISTENTLY USING A LATRINE RATHER THAN OPEN DEFECATION
1. Hearing about vaccines and the schedule	1. Learning about the need to register a birth	1. Learning about the benefits and skills of breastfeeding during pregnancy	1. Learning about the availability and value of ECE	1. Hearing about safe sanitation practices
2. Receiving a reminder for the first visit	2. Understanding the timeline and process	2. Initiating breastfeeding within the first hour after birth	2. Deciding to enrol the child	2. Accessing or constructing a functional latrine
3. Attending the first appointment on time	3. Gathering necessary documents or forms	3. Breastfeeding on demand, day and night	3. Completing registration or paperwork	3. Making the latrine convenient, clean, and private
4. Receiving and understanding return date	4. Planning when and how to go	4. Avoiding giving water, formula, or other foods	4. Preparing the child for the first day	4. Using the latrine for all defecation needs
5. Planning and preparing for the follow-up visit	5. Travelling to the registration site	5. Seeking advice when facing difficulties (e.g. pain, low milk)	5. Bringing the child consistently each morning	5. Encouraging consistent use by all household members
6. Returning for the next scheduled dose	6. Submitting the documents and completing the process	6. Sustaining exclusive breastfeeding through the 6th month	6. Continuing attendance throughout the school year	6. Maintaining the latrine regularly to ensure usability
7. Completing the full immunization schedule	7. Receiving and keeping the birth certificate			

The UNICEF Journey to Health and Immunization framework. UNICEF, Demand for health services field guide: a human-centred approach. New York: UNICEF, 2018.

d. For each journey moment, describe what the profile would ideally do if everything went as intended: this is the ideal journey. Use clear, concrete, and observable language that focuses on actions, rather than hopes or general intentions. The ideal journey should reflect what the actor does, not what they know, feel, or believe (e.g. "Caregiver brings child to clinic" rather than "Caregiver is aware of the importance of vaccines"). This should also reflect a sequence that flows logically with each moment building toward the target behaviour, assuming that key supports, services, and information are in place.

e. Describe what happens in reality. This is the real journey: the observed or reported actions that people take. For example, "Mother breastfeeds but occasionally gives water based on family advice"; "Caregiver brings the child on some days, but skips when busy or when older siblings are home". Focus on what the actor does, not on interpreting why they do it — that comes in the next step. The real journey should reflect actual behaviour based on field research, observation, and practitioner insight.

When describing the real journey, it's important to identify *drop-off points*: moments where the actor stops progressing toward the target behaviour. These are steps in the journey where actions are delayed, skipped entirely, or never completed. Drop-off points often signal high-friction moments where behavioural support is most needed. A drop-off does not necessarily mean someone has "failed"; it may be temporary (e.g. repeated delays) or more final (e.g. never returning for a service). Examples of drop-off points include:

- A caregiver brings a child for the first vaccine dose, but doesn't return for the next.
- A parent begins registration, but stops halfway due to missing paperwork.
- A mother breastfeeds exclusively for 2 months, then introduces water or food early.
- A child is enrolled in school, but stops attending regularly after a few weeks.

To highlight these moments:

- Mark them clearly in the real journey row (e.g. using a specific symbol or different colour).
- In the diagnosis step, pay special attention to these points. They are often where the most significant barriers are concentrated.
- Drop-off points are not always the final step. They can occur anywhere along the journey and still disrupt progress.

f. Next, examine each journey moment to understand why the real journey diverges from the ideal, and why the behaviour is not occurring as intended. Identify the behavioural barriers that are generating the gap between the ideal journey and the real journey, including the specific capability, behaviour, or motivation-related factors that are blocking progress at each step of the journey. This step should be based on the data already collected during the desk research and primary research, rather than speculations and assumptions. The goal is to translate qualitative and quantitative insights into clear, actionable behavioural barriers, using the COM-B model as a lens. For each barrier, use data from the research to ask:

- What is preventing the actor from doing this?
- What part of COM-B does this relate to? (Capability, Opportunity or Motivation)
- What exactly does it look like in this context?

Refer back to **Table 5. Mapping of Behavioural Drivers and Drivers in COM-B Framework** to support the diagnosis process. This table breaks down each part of the COM-B model into more specific behavioural drivers, with definitions to guide analysis.

The table is intentionally generic and should be adapted based on the specific behaviour, actor, and context. Make sure to avoid generic statements such as "low awareness" or "lack of access." Instead, describe the specific type of barrier present, and how it manifests in the real world. For example:

- Instead of noting "caregiver lacks motivation," write: "Caregiver believes one vaccine dose is enough, and does not see value in returning." (Reflective motivation: beliefs about consequences.)
- Instead of noting "clinic access is an issue," write: "Clinic only opens during market hours, when caregivers are working." (Physical opportunity: time and scheduling.)

g. Map out enablers by considering what factors are already helping the actor move toward the desired behaviour — these could already be present in the context, could be reinforced. Also consider what factors could help further move the actor towards the desired behaviour if they were strengthened. These are enablers, making the behaviour easier, more likely, or more appealing to perform. Enablers can include existing strengths, support systems, or moments of success observed during research. Like barriers, they should be described clearly, and be grounded in real data rather than assumed ideals. Return to the data and look for:

- Quotes that express positive intention or commitment
- Enabling dynamics (e.g., peer support, reminders, community mobilization)
- Aspects of the environment or routine that smooth the path to action

Enablers are not just the absence of barriers: they are active drivers of positive behaviour. Identifying them can reveal opportunities for nudges, reinforcement, or scaling what works.

h. Make sure to build a Behavioural Map and Diagnosis worksheet for each key actor involved in the behaviour. For example, if both caregivers and health workers influence a vaccination journey, build one map for the caregiver, and a separate one for the health worker. Each actor has a distinct role, decision-making process, and set of behavioural barriers, and combining them into a single map can obscure important differences.

If the research suggests significant variation within the same type of actor (for example, rural vs. urban caregivers, mothers vs. grandmothers, or younger vs. older adolescents) consider developing separate behavioural maps and diagnoses for each subgroup. This allows for a more nuanced understanding of the journey and helps to surface differences in barriers, drop-off points, or motivations that might otherwise be missed. That said, avoid over-segmenting, unless the data strongly supports meaningful differences. Focus on distinctions that are behaviourally relevant and actionable.

Here are some tips to develop a strong behavioural map and diagnosis :

- **Make it visual if needed.** Visual collaboration tools (e.g. Miro, Mural, Figma) can be used to make the mapping process more interactive and easier to navigate, especially when journeys become more complex. Colour coding barriers and enablers by COM-B domains can help to highlight patterns and clusters across the behavioural journey.
- **Consider reporting and dissemination needs.** The worksheet serves as a tool for structuring analysis, but it doesn't need to be the final output. Depending on how findings will be shared, the map can be expanded into a narrated report, visual presentation, or synthesis deck.
- **Use the draft journey as a basis for feedback and validation.** This can also be shared with subject matter experts, practitioners, or community members for feedback and validation, especially if critical gaps or assumptions are identified during analysis.

6. Prioritize key behavioural barriers and enablers

After diagnosing a broad set of behavioural barriers and enablers, it's time to narrow the focus. This sub-step helps to identify which barriers and enablers matter most, and which ones teams are best positioned to address. It ensures the design process is both grounded in insight and oriented toward impact.

Prioritization is not about discarding complexity. It's about choosing a place to begin, where our actions can remove the most friction, unlock new possibilities, and deliver measurable change.

While prioritizing different barriers, it's important to revisit the *Define* phase and understand how each barrier affects the overall problem and its broader context.

During this phase, teams move from a long list of behavioural barriers and enablers to a shortlist of 4–6 priority focus areas that will guide the *Prototype and Design* phase.

Each priority should:

- Address a clear behavioural drop-off point
- Be grounded in formative evidence
- Be actionable — something design can influence
- Offer significant potential to shift behaviour

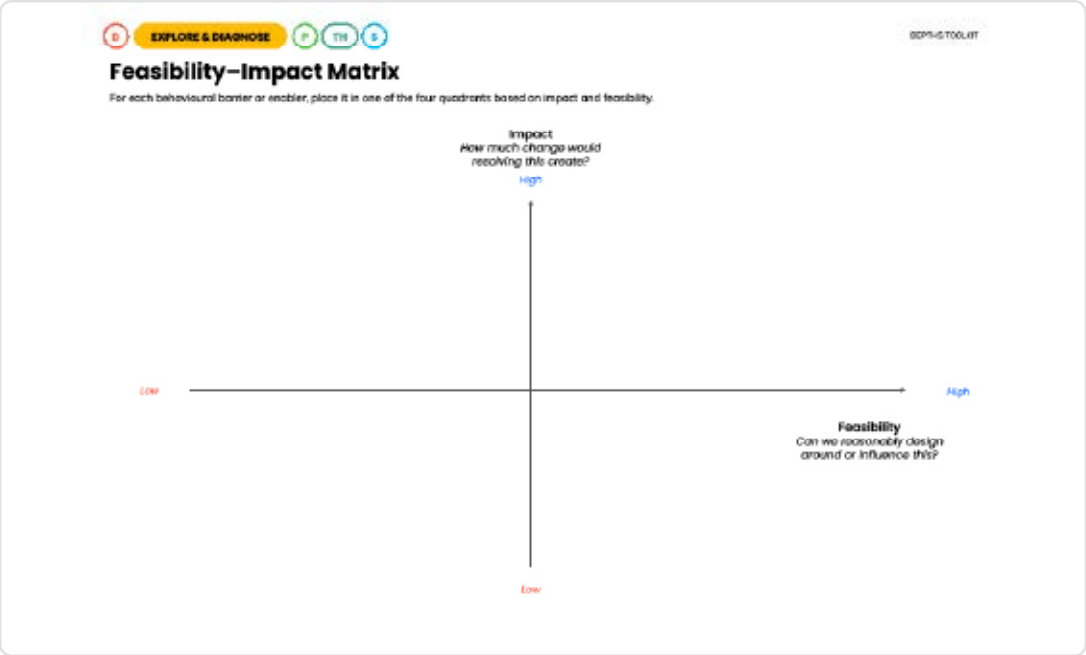
To create a shortlist of focus areas:

a. Revisit the behavioural maps and diagnoses to scan the full set of behavioural drop-off points, along with the corresponding barriers and enablers identified during diagnosis. For each drop-off point, consider:

- Which moments are most critical to enabling the full behaviour?
- Are certain barriers or enablers recurring across multiple actors or stages, suggesting systemic importance?

b. Revisit the *Define* phase's key outputs — in particular, the Stakeholder and Audience Map, the Behaviour Tree, and the optional System Map with its leverage point analysis. Reflect on the following questions:

- Does this barrier recur across different actors identified in the Stakeholder and Audience Map?
- How does this barrier affect key behaviours of different actors mapped in the Behaviour Tree?



- Where does this barrier position itself in the System Map? Does it align with an identified leverage point?
- How does this barrier compare to others in terms of equity impact? Does it exclude a specific type of audience in the Stakeholder and Audience Map?

c. Use the Feasibility–Impact Matrix to assess feasibility and impact. For each barrier or enabler, assess the following:

- **Impact:** If addressed, would this meaningfully increase the likelihood of the target behaviour occurring?
- **Feasibility:** Is it realistic to address this within the team's mandate, resources, timeline, or political/organizational constraints?

d. Map each barrier/enabler onto the matrix. This process encourages teams to design only for what is both possible and meaningful

- High Impact / High Feasibility → Prioritize for design
- High Impact / Low Feasibility → Flag for longer-term change or advocacy
- Low Impact / High Feasibility → Deprioritize, unless part of a larger enabling condition
- Low Impact / Low Feasibility → Remove from scope

e. Select 4–6 focus areas to prioritize from the top quadrants of the Feasibility-Impact Matrix. Ensure that the focus areas:

- Cover key actors and moments in the behavioural journey
- Include a mix of dynamics related to capability, opportunity, and motivation
- Reflect a coherent design opportunity — not just isolated insights

When prioritizing behaviours, look for patterns across data, as opposed to isolated anecdotes. Consider equity, and which barriers disproportionately affect marginalized groups. Teams should reflect on the question, "If this was solved, what else might become possible?"

By the end of this step, the focus has shifted from scattered field insights to a structured, behaviourally informed foundation for action. Rich behavioural profiles have been developed, and the sequence of micro-behaviours that lead to or obstruct change have been traced. Furthermore, the barriers and enablers that shape those behaviours have been diagnosed, and the most strategic areas for intervention have been prioritized. Teams now possess more than a set of findings — this is instead a grounded, human-centred understanding of how change happens, and where it gets stuck. These insights can now propel teams into the next phase: Prototype and Design, where it's time to begin translating evidence into solutions.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The Behavioural Profile and Behavioural Map and Diagnosis were not developed by the original project team. They are recreated examples based on real project data and context.

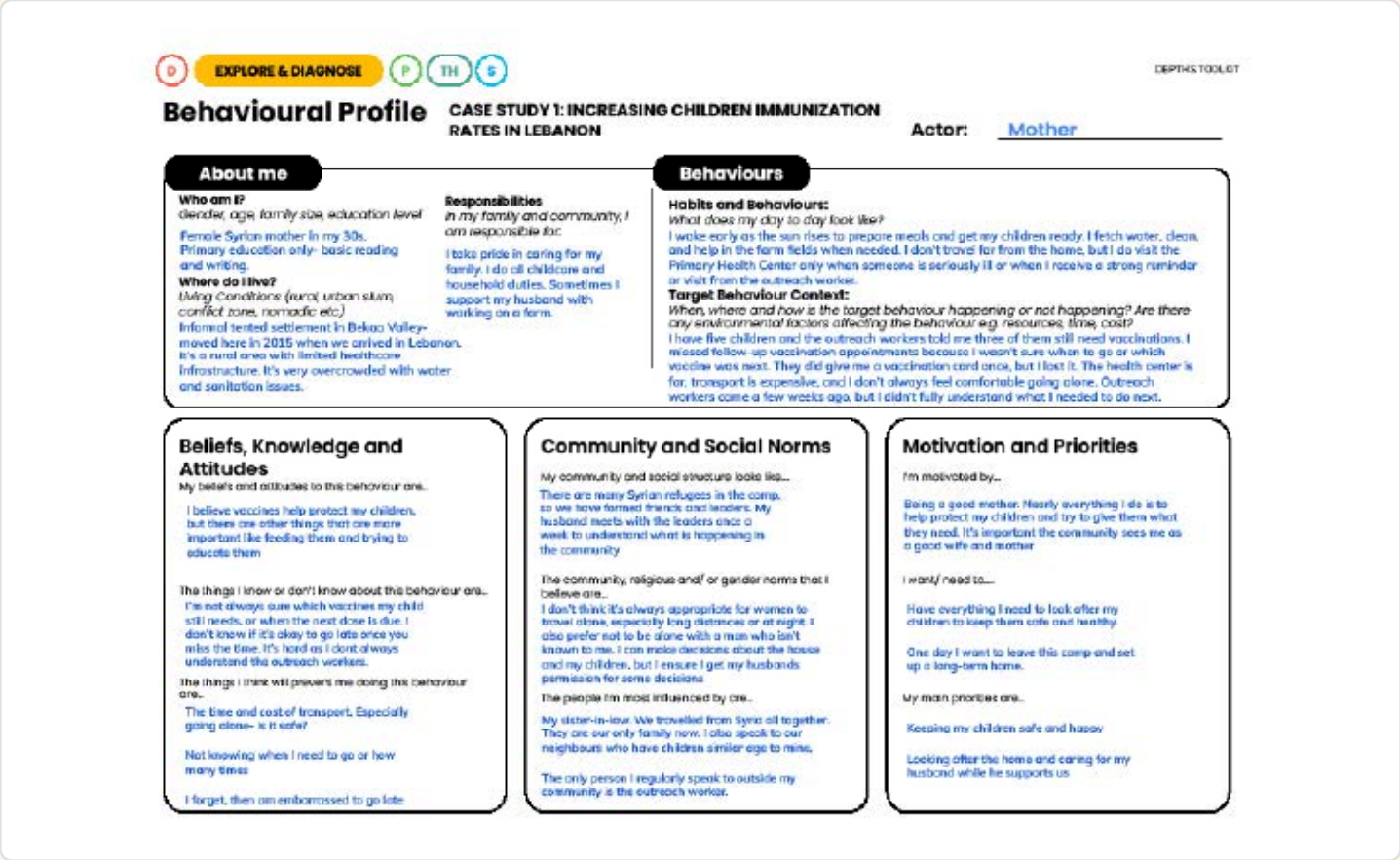
Synthesise and analyse findings

After completing the desk and primary research, the project team synthesised the data across interviews from the caregivers, health workers, outreach workers, and MoPH staff. They synthesised key data relating to the COM-B behavioural domains, along with information regarding the wider context, behaviours, beliefs, and motivations of the different participants. Based on this data, they were able to analyse and extract key insights to identify a number of personas — participants with similar contexts, behaviours, wants, and needs.

Develop a behavioural profile

The team tagged the data according to the identified personas, with one being a Syrian mother living in an informal settlement. The project team then collected all of the data they had from five research participants considered to be part of this 'persona type'. They used this data to develop a Behavioural Profile, which summarized key information relating to the persona's characteristics, overall behaviours, beliefs, motivations and norms.

This tool allowed the team to consider the wider context impacting a Syrian mother's decision-making and behaviours — beyond solely vaccination. They continued to use this Behavioural Profile throughout the project to remind themselves of who they were designing for, and what other factors may influence their decisions.



D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Behavioural Profile CASE STUDY 1: INCREASING CHILDREN IMMUNIZATION RATES IN LEBANON

Actor: Mother

About me

Who am I?
Gender, age, family size, education level
Female Syrian mother in my 30s.
Primary education only- basic reading and writing.

Where do I live?
Living Conditions (rural, urban slum, conflict zone, nomadic etc)
Informal tented settlement in Bekaa Valley- moved here in 2015 when we arrived in Lebanon.
It's a rural area with limited healthcare infrastructure. It's very overcrowded with water and sanitation issues.

Responsibilities
In my family and community, I am responsible for:
I take pride in caring for my family. I do all childcare and household duties. Sometimes I support my husband with working on a farm.

Behaviours

Habits and Behaviours:
What does my day to day look like?
I wake early as the sun rises to prepare meals and get my children ready. I fetch water, clean, and help in the farm fields when needed. I don't travel far from the home, but I do visit the Primary Health Center only when someone is seriously ill or when I receive a strong reminder or visit from the outreach worker.

Target Behaviour Context:
When, where and how is the target behaviour happening or not happening? Are there any environmental factors affecting the behaviour e.g. resources, time, cost?
I have five children and the outreach workers told me three of them still need vaccinations. I missed follow-up vaccination appointments because I wasn't sure when to go or which vaccine was next. They did give me a vaccination card once, but I lost it. The health center is far, transport is expensive, and I don't always feel comfortable going alone. Outreach workers came a few weeks ago, but I didn't fully understand what I needed to do next.

Beliefs, Knowledge and Attitudes

My beliefs and attitudes to this behaviour are...
I believe vaccines help protect my children, but there are other things that are more important like feeding them and trying to educate them

The things I know or don't know about this behaviour are...
I'm not always sure which vaccines my child still needs, or when the next dose is due. I don't know if it's okay to go late once you miss the time. It's hard as I dont always understand the outreach workers.

The things I think will prevent me doing this behaviour are...
The time and cost of transport. Especially going alone- is it safe?

Not knowing when I need to go or how many times

I forget, then am embarrassed to go late

Community and Social Norms

My community and social structure looks like...
There are many Syrian refugees in the camp, so we have formed friends and leaders. My husband meets with the leaders once a week to understand what is happening in the community

The community, religious and/ or gender norms that I believe are...
I don't think it's always appropriate for women to travel alone, especially long distances or at night. I also prefer not to be alone with a man who isn't known to me. I can make decisions about the house and my children, but I ensure I get my husbands permission for some decisions

The people I'm most influenced by are...
My sister-in-law. We travelled from Syria all together. They are our only family now. I also speak to our neighbours who have children similar age to mine.

The only person I regularly speak to outside my community is the outreach worker.

Motivation and Priorities

I'm motivated by...
Being a good mother. Nearly everything I do is to help protect my children and try to give them what they need. It's important the community sees me as a good wife and mother

I want/ need to...
Have everything I need to look after my children to keep them safe and healthy

One day I want to leave this camp and set up a long-term home.

My main priorities are...
Keeping my children safe and happy

Looking after the home and caring for my husband while he supports us

Develop a Behavioural Map and Diagnosis

This Behavioural Map and Diagnosis was not developed by the original project team. It is a recreated example based on real project data and context. Some details may not be completely accurate, but the purpose is to illustrate what a completed Behavioural Map could look like in practice.

Building on the Behavioural Profile, the project team developed a Behavioural Map and Diagnosis to trace the Syrian mother's journey toward getting her children vaccinated. Drawing from insights collected during the field research, they began by identifying key "moments" in a caregiver's immunization journey (from hearing about vaccines to completing the full series). These seven moments helped to structure the Behavioural Map and made visible the step-by-step process that caregivers navigate.

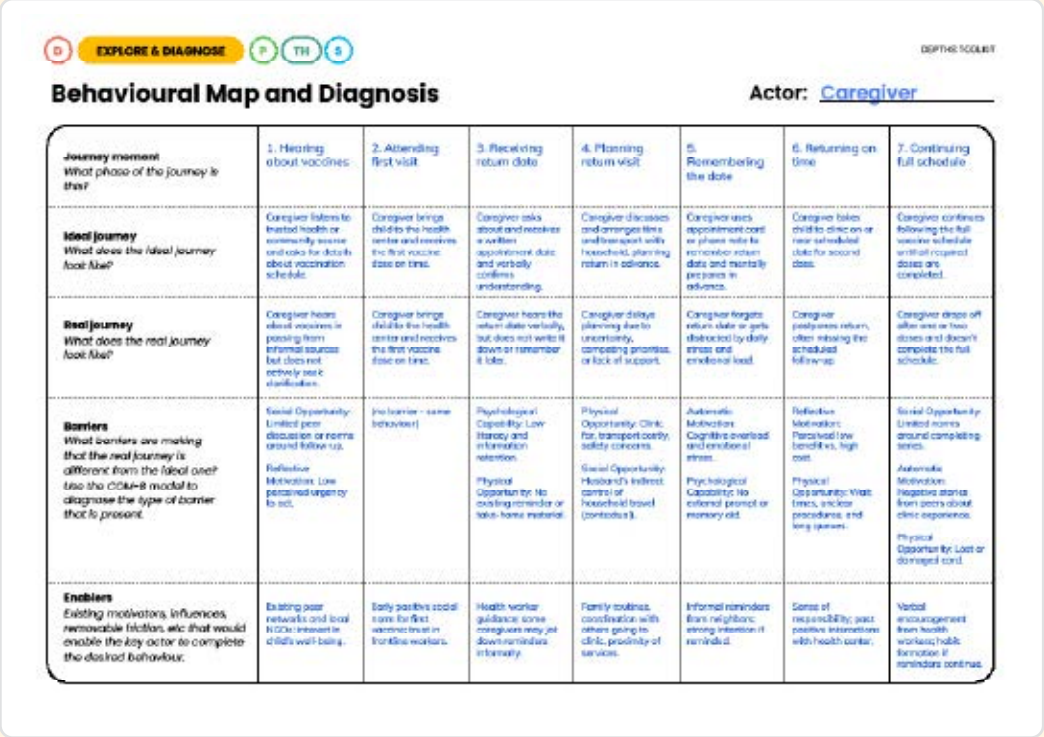
For each moment, the team then defined two factors: the ideal behaviour that would lead to full and timely vaccination, and the real behaviour caregivers were exhibiting, based on field observations and in-depth interviews. For example, at the end of their first vaccination visit, the ideal was for caregivers to ask about the date of the next vaccination and receive a note with a written date. However, the real behaviour exhibited was more uncertain. Caregivers often heard the date of the second dose vaccine verbally but had no way to record it, leading to forgetfulness later.

Next, the team diagnosed the barriers causing this gap, using the COM-B model. In the case above, the issue was one of psychological capability (difficulty retaining information) and physical opportunity (lack of take-home reminders). At the moment of remembering the date, caregivers often forgot or felt overwhelmed, which the team coded as automatic motivation, shaped by cognitive overload and emotional stress.

However, not all behaviours diverged. For example, in the second journey moment (i.e., attending the first vaccine visit) caregivers were largely consistent with the ideal. This revealed that the real bottlenecks began after the first visit, reinforcing the importance of focusing interventions on follow-up behaviours rather than initial awareness.

Finally, the team identified existing enablers at each step. These included verbal encouragement from health workers, positive early experiences with vaccines, and informal reminders from peers. This process helped the team to identify which micro-behaviours mattered most, why they were breaking down, and where low-cost nudges could support change. The Behavioural Map became a foundational diagnostic tool to guide ideation and prototype development.

While this example focuses on Syrian caregivers, the broader research also revealed distinct barriers among Lebanese families. These included a lack of trust in the quality of vaccines and services offered, and a belief that they were entitled to better care than what was provided through programmes aimed at non-Lebanese populations.



D EXPLORE & DIAGNOSE P TH S

DEPTHS TOOLKIT

Behavioural Map and Diagnosis

Actor: Caregiver

Journey moment *What phase of the journey is this?*	1. Hearing about vaccines	2. Attending first visit	3. Receiving return date	4. Planning return visit	5. Remembering the date	6. Returning on time	7. Continuing full schedule
Ideal journey *What does the ideal journey look like?*	Caregiver listens to trusted health or community source and asks for details about vaccination schedule.	Caregiver brings child to the health center and receives the first vaccine dose on time.	Caregiver asks about and receives a written appointment date and verbally confirms understanding.	Caregiver discusses and arranges time and transport with household, planning return in advance.	Caregiver uses appointment card or phone note to remember return date and mentally prepares in advance.	Caregiver takes child to clinic on or near scheduled date for second dose.	Caregiver continues following the full vaccine schedule until all required doses are completed.
Real journey *What does the real journey look like?*	Caregiver hears about vaccines in passing from informal sources but does not actively seek clarification.	Caregiver brings child to the health center and receives the first vaccine dose on time.	Caregiver hears the return date verbally, but does not write it down or remember it later.	Caregiver delays planning due to uncertainty, competing priorities, or lack of support.	Caregiver forgets return date or gets distracted by daily stress and emotional load.	Caregiver postpones return, often missing the scheduled follow-up.	Caregiver drops off after one or two doses and doesn't complete the full schedule.
Barriers *What barriers are making that the real journey is different from the ideal one? Use the COM-B model to diagnose the type of barrier that is present.*	Social Opportunity: Limited peer discussion or norms around follow-up. Reflective Motivation: Low perceived urgency to act.	(no barrier – same behaviour)	Psychological Capability: Low literacy and information retention. Physical Opportunity: No cooking reminder or take-home material.	Physical Opportunity: Clinic, for transport costs; safety concerns. Social Opportunity: Husband's indirect control of household travel (contextual).	Automatic Motivation: Cognitive overload and emotional stress. Psychological Capability: No external prompt or memory aid.	Reflective Motivation: Perceived low benefits, high cost. Physical Opportunity: Wait times, unclear procedures, and long queues.	Social Opportunity: Limited norms around completing series. Automatic Motivation: Negative stories from peers about clinic experience. Physical Opportunity: Lost or damaged card.
Enablers *Existing motivators, influences, removable friction, etc that would enable the key actor to complete the desired behaviour.*	Existing peer networks and local NGOs interested in child's well-being.	Early positive social norm for first vaccine; trust in frontline workers.	Health worker guidance; some caregivers may jot down reminders informally.	Family routines, coordination with others going to clinic, proximity of services.	Informal reminders from neighbors; strong intention if reminded.	Sense of responsibility; past positive interactions with health center.	Verbal encouragement from health workers; habit formation if reminders continue.

Apply a Feasibility–Impact Matrix

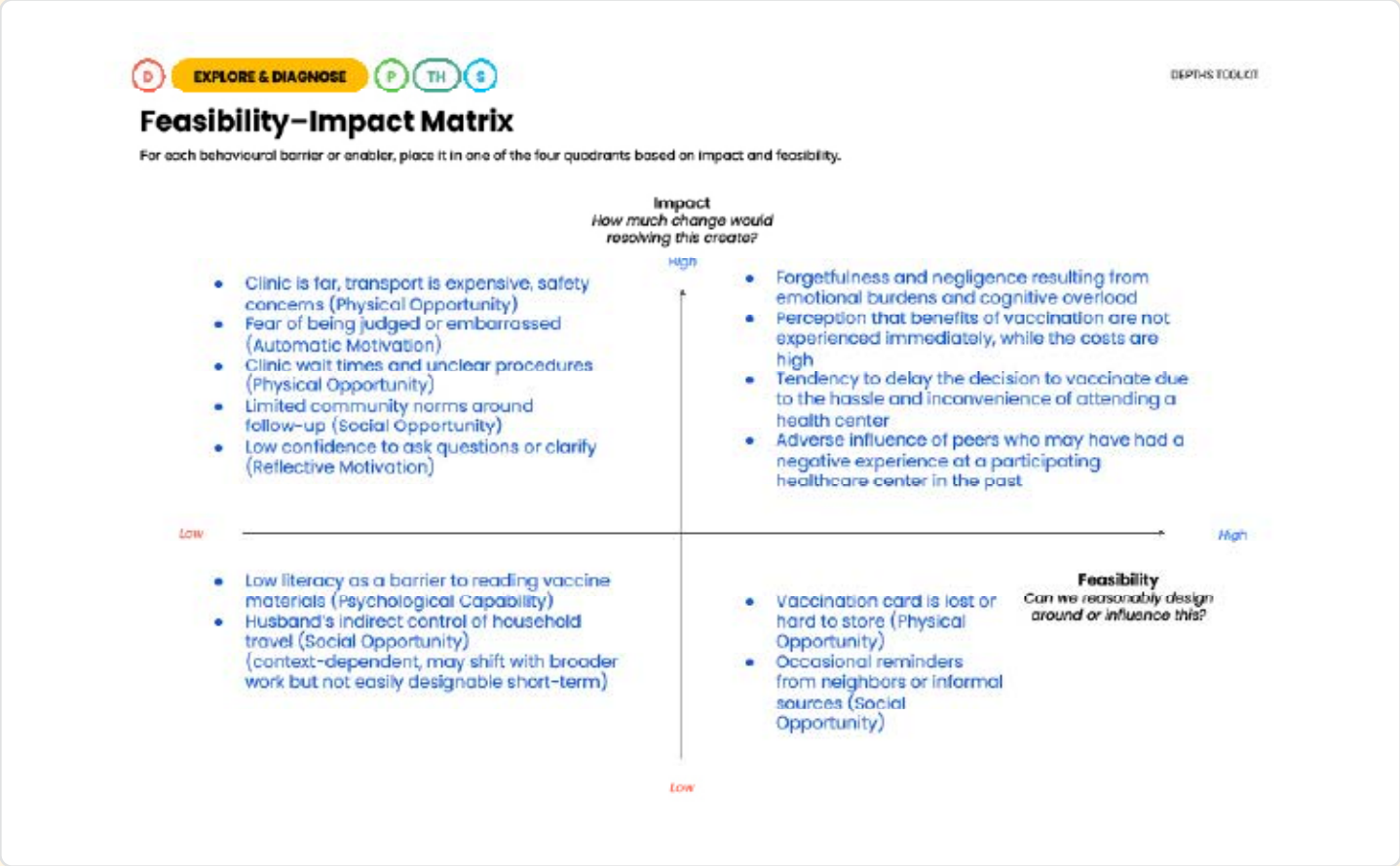
This Feasibility-Impact Matrix was not developed by the original project team. It is a recreated example based on real project data and context. Some details may not be completely accurate, but the purpose is to illustrate what a completed Feasibility-Impact Matrix could look like in practice.

After completing the Behavioural Map and Diagnosis, the project team used the Feasibility-Impact Matrix to prioritize which barriers and enablers should be addressed in the design phase. Each factor was assessed based on its potential impact on the caregiver's ability to complete the vaccination journey, and the feasibility of influencing it through behavioural design. The barriers were also considered against the System Map and the Behaviour Tree that were developed in the previous Define phase. This helped the team to understand the role each barrier played in the broader context of the problem.

This process aided the team in their transition from identifying behavioural drivers to selecting practical points of intervention. While certain challenges — such as the long distance to the clinic — were recognized as important, they were considered less feasible to address directly within the scope of the project. In contrast, the team prioritized barriers that were both highly influential and actionable. These included the following:

- forgetfulness and inaction linked to emotional strain and cognitive overload
- the perception that vaccination benefits are delayed while the effort and side effects are immediate
- a tendency to postpone vaccination due to the inconvenience of attending health services
- the discouraging influence of peers who had previously experienced poor treatment at healthcare centers

By organizing these insights within the matrix, the team identified a focused set of behavioural challenges to guide the design of interventions. This ensured that their efforts were both strategic and responsive to the daily experiences of caregivers.



Final checklist for *Explore & Diagnose*

- [ ] Initial Desk Research
- [ ] Primary Research Objectives and Questions
- [ ] Primary Research Plan and Protocol
- [ ] Behavioural Map and Diagnosis
- [ ] Feasibility-Impact Matrix

Optional:

- [ ] Behavioural Profile

Learn more

This field guide is designed with practical tools, frameworks, and methodologies for teams to apply behavioural science to a range of real-world challenges. As behavioural science draws from multiple disciplines, including human-centred design, experimental economics, and systems thinking, we've curated a selection of approaches that reflect this diversity. Behavioural science is a vast and constantly growing field, and it would be impossible to capture every tool or perspective in a single guide — that's why we've included this section for those who are curious to go deeper. The following section offers additional resources to explore specific topics introduced in the guide, offering an option to continue a self-paced learning journey.

"I want to understand the foundations of qualitative research and when to use it."

If you're new to qualitative research or want a clear, structured introduction, Allison Hurst's Introduction to Qualitative Research Methods is a fantastic place to start. This free textbook walks through everything from research questions to analysis, with many real-world examples to help bring concepts to life.

To deepen your understanding of when qualitative methods are appropriate and how to defend their rigor, turn to Hammarberg et al.'s article (2016) in Human Reproduction. The article lays out what makes qualitative research credible, and offers practical advice for explaining methodological choices to funders, reviewers, or skeptics who may be unfamiliar with the approach.

"I want to learn how to identify high quality evidence."

One widely used resource and framework used to assess the quality of evidence is the GRADE approach, commonly found in meta-analyses and systematic reviews. Essentially, GRADE provides a transparent and structured method to rate our confidence in the available evidence. It categorizes evidence quality into different levels based on factors that can downgrade or upgrade our confidence, such as risk of bias, consistency, directness, and precision. This helps reviewers to evaluate not only what the evidence says, but how much trust they can place in those findings.

"I want to learn how to decide on sample size for qualitative studies."

Forget one-size-fits-all rules. Instead, explore the concept of "information power" introduced by Malterud et al. (2016). This idea helps researchers to focus on the value of information each participant brings to a study, rather than arbitrary notions of saturation — providing a more strategic and meaningful way to plan your sample.

"I want to adopt a decolonial approach to my qualitative research."

The Center for Critical Qualitative Health Research (CCQHR) at the University of Toronto offers a wealth of advanced materials. Their website includes curated reading lists on Indigenous and decolonial methods, health anthologies in Spanish and Portuguese, and recorded lectures from leading thinkers. All are freely available online and on their YouTube channel.

A seminal resource in this field is *Decolonizing Methodologies: Research and Indigenous Peoples* (2nd ed., 2012) by Linda Tuhiwai Smith, which critically examines how research practices intersect with colonial histories and offers practical approaches for decolonising research. The textbook is available for purchase, or can be accessed for free here.

"I want access to journals and ongoing learning opportunities."

To stay up to date with emerging ideas and methodologies, consider exploring two open-access journals: the International Journal of Qualitative Methods (IJQM) publishes high-quality articles that push methodological boundaries, while Forum: Qualitative Social Research (FQS) has been sharing innovative work in English, German, and Spanish for over two decades.

"I want guidance on how to design and analyse qualitative research."

If you're ready to take a deeper dive into designing your own qualitative research, these textbooks are essential:

- FHI 360 – Qualitative Research Field Guide, which includes practical steps for introducing Qualitative Research Methods.
- Creswell & Poth's Qualitative Inquiry and Research Design guides you through five major approaches including narrative and grounded theory, with detailed examples.
- Green & Thorogood's Qualitative Methods for Health Research is especially relevant for those working in public health or clinical settings.
- Maxwell's Qualitative Research Design emphasizes the importance of coherence and interactivity in research design, with tools to help build strong proposals.

"I want to learn more about how to use different quantitative approaches for my research."

Exploratory quantitative research often begins with surveys, descriptive statistics, and cross-tabulations to examine relationships between behaviours, beliefs, and context. Start by learning how to frame good research questions, choose appropriate indicators, and structure survey instruments that generate useful, analyzable data.

The DIME Analytics Data Handbook is an excellent step-by-step guide to managing data workflows (from planning to collection, coding, and storage) tailored for teams working in real-world development settings. For a more structured academic foundation, the Springer textbook on Quantitative Methods for the Social Sciences provides accessible explanations of descriptive statistics, variable types, and data visualization techniques. For those working in health contexts, Mixed Methods in Health Sciences Research: A Practical Primer by Curry and Nunez-Smith offers guidance on how to thoughtfully combine quantitative and qualitative approaches for real-world research questions.

"I want to learn more about different behavioural frameworks to guide my research and analyse data."

For those aiming to learn more about different behavioural frameworks that could be used for both guiding the primary research, as well as analyzing qualitative data, here are some additional resources to learn more about frameworks that have been mentioned in this chapter:

- The World Health Organization's Behavioural and Social Drivers of Vaccination (BeSD) Framework provides a structured tool for diagnosing and addressing the drivers of vaccine uptake. It categorizes influences into four domains: what people think and feel, social processes, motivation, and practical issues — making it particularly suited for immunization-related research and planning.
- The Socioecological Model which was first developed by Urie Bronfenbrenner is a widely used framework that focuses on different networks surrounding a person. UNICEF has a tailored version for children.
- The Behavioural Drivers Model (BDM) offers a flexible, context-sensitive framework that allows teams to explore the cognitive, social, emotional, and environmental factors that shape human behaviour. It is particularly useful in formative research and in identifying opportunities for targeted interventions.
- For those seeking a simpler model grounded in behavioural psychology, Fogg's B-MAP (Behaviour–Motivation–Ability–Prompt) model provides an accessible lens through which to understand and influence behaviour.

Resources

1. Ajzen, Icek. "The Theory of Planned behaviour: Frequently Asked Questions." Human behaviour and Emerging Technologies 2, no. 4 (2020): 314–324.
2. Bakrania, Shivit, ed. Methodological Briefs on Evidence Synthesis. UNICEF Office of Research – Innocenti, 2020. <https://www.unicef.org/innocenti/documents/methodological-briefs-evidence-synthesis>.
3. Bjärkefur, Kristoffer, Luiza Cardoso de Andrade, Benjamin Daniels, and Maria Ruth Jones. Development Research in Practice: The DIME Analytics Data Handbook. Washington, DC: World Bank, 2021.
4. Braun, Virginia, and Victoria Clarke. "Using Thematic Analysis in Psychology." Qualitative Research in Psychology 3 (2006): 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
5. Busara Center for Behavioural Economics. Barriers and Levers Framework. Nairobi: Busara.
6. Busara Center for Behavioural Economics. Conducting Rigorous Qualitative Research. Nairobi: Busara.
7. Busara Center for Behavioural Economics. Mixed Methods Research Toolkit. Nairobi: Busara.
8. Center for Advanced Hindsight. behavioural Mapping Cheat Sheet. Durham: Duke University.
9. Center for Critical Qualitative Health Research, University of Toronto. "Resources and Publications." <https://ccqhr.utoronto.ca/resources/publications>.
10. Copper, S., and K. Perry. "Survey Design." J-PAL, 2023. <https://www.povertyactionlab.org/resource/survey-design>.
11. Creswell, John W. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 4th ed. Thousand Oaks, CA: Sage Publications, 2014.
12. Decision Lab. COM-B Framework. Available at <https://www.thedlab.com>.
13. Empatika. Social Determinants Influencing Access to Malaria Services: A Formative Study in NTT, Papua and West Papua. UNICEF Indonesia, 2021. https://www.unicef.org/indonesia/health/reports/social-determinants-influencing-access-malaria-services.
14. Fogg, B. J. Fogg behaviour Model: B=MAP. Stanford: Behaviour Design Lab.
15. Forum: Qualitative Social Research (FQS). <https://www.qualitative-research.net>.
16. Green, Judith, and Nicki Thorogood. Qualitative Methods for Health Research. 4th ed. London: Sage Publications, 2018.
17. Hammarberg, Karin, Maggie Kirkman, and Shelley de Lacey. "Qualitative Research Methods: When to Use Them and How to Judge Them." Human Reproduction 31, no. 3 (2016): 498–501. <https://doi.org/10.1093/humrep/dev334>.
18. Hurst, Allison. Introduction to Qualitative Research Methods. Corvallis, OR: Oregon State University, 2023. <https://open.oregonstate.education/qualresearchmethods/>.
19. Interaction Design Foundation. behavioural Personas. <https://www.interaction-design.org>.
20. International Institute for Environment and Development. Stakeholder Power Analysis. London: IIED.
21. International Journal of Qualitative Methods (IJQM). <https://journals.sagepub.com/home/IJQ>.
22. Jackson II, Ronald, David Drummond, and Sakile Camara. "What Is Qualitative Research?" Qualitative Research Reports in Communication 8 (2007): 21–28. <https://doi.org/10.1080/17459430701617879>.
23. J-PAL. IRB Proposals. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
24. J-PAL. Research Resources. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
25. J-PAL. Survey Design. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.

26. Kauschanski, Alexander. The Power of Youth: Instruments and Case Studies for Effective Youth Participation. Desk research, UNICEF, January 2020. <https://www.unicef.org/ukraine/en/reports/power-youth-instruments-and-case-studies-effective-youth-participation>

27. Malterud, Kirsti, Volkert D. Siersma, and Ann Dorrit Guassora. "Sample Size in Qualitative Interview Studies: Guided by Information Power." Qualitative Health Research 26, no. 13 (2016): 1753–1760. <https://shorturl.at/7p1bt>.

28. Maxwell, Joseph A. Qualitative Research Design: An Interactive Approach. 3rd ed. Thousand Oaks, CA: Sage Publications, 2013.

29. Michie, Susan, Maartje M. van Stralen, and Robert West. "The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions." Implementation Science 6 (2011): 42. <https://doi.org/10.1186/1748-5908-6-42>.

30. Ng, Candy F. "Behavioural Mapping and Tracing." In Research Methods for Environmental Psychology, edited by Robert Gifford, 29–52. Wiley, 2016.

31. Patton, Michael Quinn. Qualitative Research & Evaluation Methods. 3rd ed. Thousand Oaks, CA: Sage Publications, 2002.

32. SAGO. Sampling for Formative Qualitative Research. <https://sago.com>.

33. Schünemann, Holger J., Jan Brożek, Gordon Guyatt, and Andrew Oxman, eds. GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations. <https://gdt.gradepro.org/app/handbook/handbook.html>.

34. Scottish Government. The ISM User Guide: Influencing Behaviours through the Individual, Social and Material Context. Edinburgh: Scottish Government.

35. The Behavioural Insights Team. Explore Field Guide: A Practical Tool to Map and Unpack Behaviour. 2022. <https://www.bi.team/wp-content/uploads/2022/11/BIT-Explore-Report-1.pdf>.

36. The SAGE Handbook of Qualitative Data Analysis. London: SAGE Publications.

37. The Uncertainty Project. Mini-Delphi Method. Available at <https://uncertaintyproject.org>.

38. UK Aid. Political Economy Analysis: A Beginner's Guide. London: DFID.

39. UK Government (Public Health England). How to note: Assessing the Strength of Evidence (Guidance Note). March 2014. Accessed August 27, 2025. <https://assets.publishing.service.gov.uk/media/5a7c95eeed915d6969f45eb5/HTN-strength-evidence-march2014.pdf>.

40. UNESCO Institute for Statistics. Guide to Conducting an R&D Survey. Paris: UNESCO.

41. UNICEF. Behavioural Drivers Model (BDM). New York: United Nations Children's Fund.

42. UNICEF SBC Guidance. Building an SBC Strategy. UNICEF, n.d. <https://www.sbcguidance.org/create/building-sbc-strategy>.

43. UNICEF. Ethical Standards in Research. New York: United Nations Children's Fund.

44. UNICEF. Global Multisectoral Operational Framework for the Use of behavioural and Social Sciences in Programming. New York: United Nations Children's Fund, 2022. <https://www.unicef.org/media/135011/file/Global%20multisectoral%20operational%20framework.pdf>.

45. University of Warwick. Literature Review, Questions, and Methods. Coventry: University of Warwick.

46. vLab. Overview of Qualitative Methods. Available at <https://vlab.org>.

47. World Bank. "LSMS Practical Guide to Fieldwork Training." World Bank Blogs. <https://blogs.worldbank.org>.

48. World Health Organization and UNICEF. Behavioural and Social Drivers of Vaccination: Tools and Practical Guidance for Achieving High and Equitable Immunization Coverage. Geneva: WHO, 2022. <https://www.who.int/publications/i/item/9789240049680>.

Prototype Designs

Welcome to Prototype Designs!

In this phase, insight becomes action. Here, the objective is to develop and prototype potential interventions to change the target behaviour(s) or micro-behaviour(s) — by addressing the prioritised barriers and leveraging the enablers identified in the previous phase, *Explore & Diagnose*.

Grounded in evidence and behavioural theory, these prototypes are brought to life through creative design and co-creation, then refined through iterative feedback. This process allows for identifying the most promising solutions for piloting and real-world testing.

Why Prototype Designs?

It's easy to fall back on familiar solutions, especially those that have worked before in the past, or in other settings. While this approach can reduce uncertainty, it can also limit creativity and opportunities to design more effective, locally relevant, or cost-efficient interventions. That's why it's important to deliberately explore new possibilities wherever possible.

The *Prototype Designs* phase focuses on generating and testing practical responses towards identified behavioural challenges. This stage encourages the development of a wide range of potential interventions rooted in behavioural science. The most promising ideas are then selected for prototyping and tested with real users. This ensures that solutions are both innovative and tailored to the specific barriers uncovered during research.

How can we Prototype Designs?

In this third phase of the *DEPTHS* methodology, there are four steps to follow through the behavioural intervention design process.

Each step includes specific tools, guidance on why these tools matter, and how to use them. A **case study on increasing childhood vaccination uptake in Lebanon** will exemplify how the tools can be applied in practice.

Before starting to *Prototype Designs*, make sure to complete all the steps of the *Define* and *Explore & Diagnose phases*.

Summary of the Explore & Diagnose phase

1	2	3	4
Ideate interventions	**Prioritize interventions**	**Prototype, test, and iterate interventions**	**Theory of Change**
Translate key insights from the Explore & Diagnose phase into actionable prompts. Use these questions and different frameworks to brainstorm a wide range of behaviourally-informed ideas.	Apply clear criteria such as feasibility, impact, scalability, and ethics to identify the most promising idea(s) to prototype.	Bring the selected idea(s) to life through prototyping to be tested with users to gather early feedback.	Refine the intervention based on user testing, and map out how and why it is expected to lead to the desired behavioural outcomes, and impact the outcome statement.
TOOLS	**TOOLS**	**TOOLS**	**TOOLS**
Ideation Prompts	Idea Prioritization	Prototype Outline	Theory of Change
Idea Generation	Ethics Checklist	Prototype Discussion Guide	
		Prototype Synthesis	

Common pitfalls

There are common traps that can derail behaviourally-informed work. Throughout the *Prototype Designs* phase, it's worth keeping these possible issues in mind:

- **Relying on familiar approaches instead of true prototyping.** In behavioural science, it is easy to default to well-known interventions like reminders, generic social norm messages, or communication campaigns built on KAP (knowledge, attitudes, practices) surveys. While these approaches can generate useful insights or raise visibility, they might not address the deeper behavioural drivers uncovered during research. Effective prototypes must be rooted in real barriers and enablers, such as caregiver routines, health system realities, or social norms, and move beyond messaging toward meaningful behaviour change.
- **Treating ideation as a blank canvas.** While creativity is essential, ideation without structured brainstorming can drift into impractical or misaligned ideas. Interventions are more likely to succeed when they are anchored in behavioural evidence and respond to specific, diagnosed barriers. For example, designing a digital app may appear innovative, but if internet access is limited or device ownership is low, the idea won't be effective. Similarly, proposing group education sessions with caregivers may miss the mark if norms discourage public discussion of decisions regarding children.
- **Finding the balance.** The key is to strike a balance: ground interventions in behavioural evidence and local insight, while staying open to creative, context-specific solutions. Structure doesn't stifle innovation, it focuses it. Similarly, adaptation doesn't mean discarding proven tools — it means making them fit the realities of the people to be supported.
- **Skipping small-scale testing.** Rushing into the field without pausing for early feedback or iterating the input of those involved can result in design decisions that are difficult to revisit later. Prototyping helps to surface confusion, friction, or unintended effects before investing further in the process. Testing early, often, and when possible, inexpensively, helps to improve ideas while they're still flexible.
- **Treating equity as a simple checklist.** Even well-intentioned interventions can cause harm if they are not designed and tested with care. Behaviourally informed ideas often interact with people's emotions, identities, and their experiences, and this comes with responsibility. Applying equity as a lens throughout the ideation process means continuously asking: Who might be left out? Who might feel judged or uncomfortable? Could this intervention unintentionally reinforce stigma, or burden vulnerable groups? Equity is central to ensuring that the solutions being developed are fair, inclusive, and truly beneficial to the communities they aim to serve.

CASE STUDY:

Increasing childhood vaccination in Lebanon

In Lebanon, rising poverty levels and the influx of over a million refugees created significant strain on the country's immunization system. Despite the launch of an accelerated immunization programme aimed at expanding access to free vaccines through community outreach, uptake remained low in several districts — particularly among refugee and low-income populations. In response, a multidisciplinary team from Nudge Lebanon, UNICEF, and the Ministry of Public Health (MoPH) came together to better understand the behavioural and contextual factors contributing to low return rates for childhood vaccination.

As part of the *Define* phase, the team used available programme data to map vaccination coverage across regions and identified districts with the lowest uptake. They focused on vulnerable populations, including informal tented settlements and under-resourced communities, and worked with local partners to understand the broader system surrounding childhood immunization. Through stakeholder mapping and root cause exercises, they identified gaps — not only in access, but also caregiver follow-through and clinic readiness.

In the *Explore & Diagnose phase*, the team conducted in-depth fieldwork using behavioural science tools. Through interviews with caregivers, healthcare providers, and outreach workers — along with the observation of household visits — they mapped the caregiver journey and uncovered critical behavioural drop-off points. Caregivers were not necessarily opposed to vaccination, but daily survival needs often made returning to the clinic difficult. Barriers included forgetting appointments, cognitive overload, and unclear instructions. These insights helped the team pinpoint where behaviour was breaking down.

Having identified the key behavioural barriers preventing caregivers from returning on time for their child's next vaccination appointment, the team generated structured "how might we" questions and brainstormed a range of creative and behaviourally informed solutions. After applying prioritization and ethical criteria, one idea stood out: a simple paper-based appointment card designed to remind caregivers of their child's next visit in a clear, tangible way.

The intervention was developed as a low-fidelity prototype and tested through informal walkthroughs with caregivers in local clinics. Feedback was gathered using a structured guide, revealing what elements worked and what needed improvement. Based on this input, the team refined the design before selecting it for implementation and evaluation.

***Note:** While this is a real project that closely followed a very similar process to DEPTHS, there were a few tools from the toolkit that the project team did not apply during implementation. In those cases, we've gone back and retrospectively applied the tools using real project data to illustrate how they might have looked if they had been used at the time.*

STEP 1:
Ideate and co-create interventions

In this step:

This step bridges insights uncovered in the Explore & Diagnose phase with the development of concrete intervention ideas, using two main tools.

The Ideation Prompts worksheet helps to formulate "how might we questions that draw on evidence gathered from both fieldwork and desk research. These serve as prompts to translate key findings into actionable opportunities, helping to channel creative thinking toward practical, testable intervention ideas aligned with users' lived experiences.

Associated Tools:

- Ideation Prompts
- Idea Generation

The Idea Generation worksheet supports creative and strategic thinking by helping to generate innovative, and meaningful ideas. This tool introduces brainstorming exercises, along with guidance to select the exercises best suited to the specific context.

Why it matters:

This step helps to avoid jumping to solutions too quickly. Instead, combining behavioural insights with lived experience and creative exploration increases the chances of generating interventions that are innovative, feasible, acceptable, and effective in driving real-world change.

How to do it:

1. Turn behavioural insights into idea prompts through "how might we" questions

Use the Idea Prompts worksheet to translate key findings from the first two phases into actionable, opportunity-focused "how might we" questions that can spark creative, locally relevant solutions:

a. Start by revisiting key insights from the Define and Explore & Diagnose phases, to better recall the nuances of the context at hand. In particular:

- **The desired outcomes and leverage points** (Project Canvas and System Map)
- **Past interventions covered by the desk research:** Any relevant efforts or programmes identified through the initial desk research, which may offer lessons or starting points for intervention design.
- **Actors, behaviours, and micro-behaviours:** The specific people we aim to support and the behaviours we want to shift (Behavioural Profiles[1] and Behavioural Map and Diagnosis)
- **Known opportunities:** Drop-off points, behavioural barriers, and enabling factors identified during the Behavioural Map and Diagnosis and prioritized using the Feasibility–Impact Matrix.
- **Data points and quotes:** Quotes reveal how people think, feel, and talk about an issue. For example, a caregiver might say, "*I was told to keep taking the TB medicine, but it makes me dizzy and I can't work when I use it.*" This highlights barriers linked to side effects and livelihood pressures that surveys may miss. Data, meanwhile, validate and quantify these patterns. For instance, clinic records showing frequent treatment interruptions confirm the challenge described in interviews.

b. Using the Idea Prompts worksheet, write down the key inputs needed to formulate the questions. These include four key components:

- **Population of interest:** Start by identifying the group or individual whose behaviour the intervention aims to influence, using outputs from previous phases. Populations can be defined by role (e.g. caregivers, healthcare workers, teachers), life stage (e.g. new parents, adolescents), or context (e.g. displaced populations, rural residents). Relevant demographic factors such as age, gender, or socio-economic status may also be included.
- **Micro-behaviour to change:** Identify the specific micro-behaviour the population of interest is expected to adopt or shift (as identified in the Behavioural Map and Diagnosis). For example, if the broader behaviour change goal is caregivers completing their children's full routine immunization

1 If the Behavioural Profile was not completed, this can be drawn from the Project Canvas from the Define phase

schedules, a target micro-behaviour could be returning for a second vaccine dose. Other examples of micro-behaviours include:

- If the broader behaviour change goal is exclusive breastfeeding for the first six months, the micro-behaviour could be initiating breastfeeding within the first hour after birth.
- If the broader behaviour change goal is ensuring all children are enrolled in school on time, the micro-behaviour could be completing school registration paperwork.

- **Barriers and Enablers:** List the key factors that help or hinder the micro-behaviour of interest, drawing from the Behavioural Map and Diagnosis and the prioritized barriers identified in the Feasibility–Impact Matrix. These barriers or enablers, aligned with the COM-B model, should relate to capability (e.g. capability: lack of knowledge or skills), motivation (e.g. beliefs and values), and opportunity (e.g. access to services or lack of social support). Refer to the Explore & Diagnose phase for more information about how to use COM-B to diagnose barriers and enablers.
- **Additional information:** Use quotes, data, or contextual details from desk research or the Behavioural Map and Diagnosis to ground this step in real experiences. These details can inspire ideas for the "how might we" prompt and keep the process connected to what people have actually said, done, or felt, instead of relying on assumptions.

c. Next, take these key components and turn them into well-structured "how might we" questions. To generate these questions, the following structure is recommended:

How might we + C barrier addressed/enabler leveraged + for A population of interest + so that B micro-behaviour will change

-
-
-
-
-
-
-

How might we + (barrier addressed or enabler leveraged) + for (population of interest) + so that (micro-behaviour will change)

This structure is only a guide. The wording can be adapted, as long as the essential elements are included. Examples of "how might we" questions across themes are:

- How might we reduce the difficulties of caregivers in remote areas reaching the clinic, so that they complete their child's vaccination schedule?
- How might we increase the perceived value of school attendance for adolescent girls in pastoralist communities, so that they attend school regularly?
- How might we strengthen trust in reporting systems for women living in informal settlements, so that they report incidents of abuse when they occur?
- How might we simplify the birth registration process for first-time parents, so that they register their child within the first month?
- How might we reduce fear of judgment from peers for adolescent boys, so that they are more willing to seek mental health support?

It's also important to remember that there is no single "correct" version of a "how might we" question. In fact, it's useful to generate several options (at least four) to explore different angles. For example:

- One version might address a barrier (e.g. reducing fear)
- Another could focus on an enabler (e.g. strengthening peer support)
- Some might centre on the primary user (e.g. caregivers), while others target key influencers (e.g. health workers)

2. Generate intervention ideas

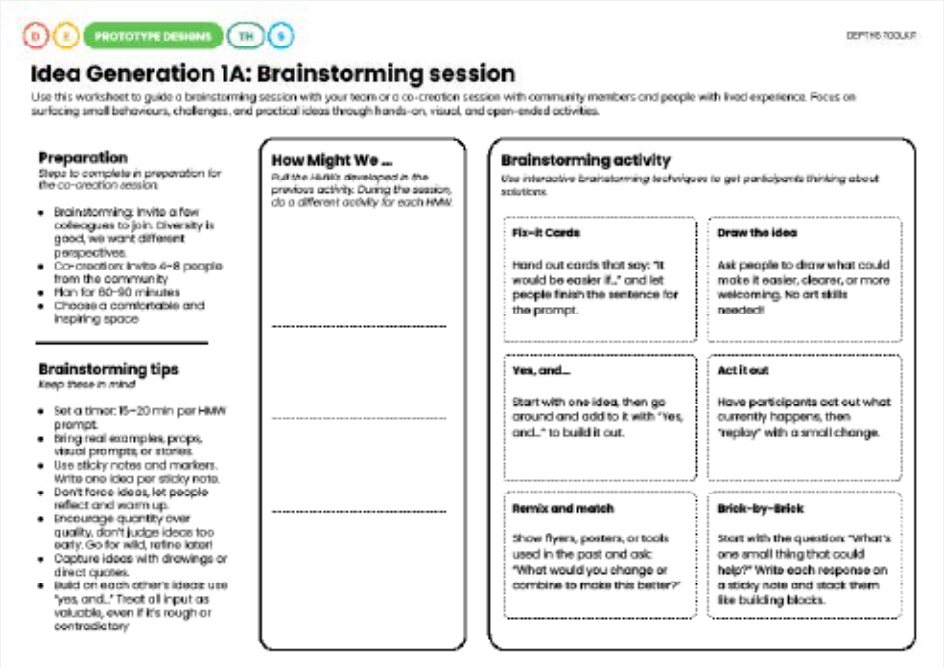
Generating strong intervention ideas isn't about finding a single "right" answer, it's a creative and iterative process without a one-size-fits-all approach. The Idea Generation worksheet provides a set of structured, manageable activities to guide this process.

Whenever possible, it is strongly recommended to hold in-person brainstorming sessions with community members, internal stakeholders, or project team members. These sessions can foster a richer discussion, stronger rapport, and more dynamic collaboration. When in-person engagement is not feasible, virtual workshops can be conducted using collaborative tools like Miro or Mural to simulate an interactive environment. If planning and facilitating remote workshops, it may be helpful to engage expert facilitators who specialize in virtual idea generation.

2.1. In-person brainstorming sessions with the community

Start by hosting brainstorming sessions to generate early intervention ideas. This step ensures that diverse perspectives shape the solutions from the outset, especially those most affected by the behavioural challenge. It's recommended to sequence brainstorming in two stages:

- **Stage 1: Internal team session.** Begin with the core project team to review behavioural insights, define the behavioural challenge, and generate ideas. Use tools like the "how might we" worksheet and revisit the Behavioural Map and Diagnosis to anchor brainstorming in evidence.
- **Stage 2: Co-creation with communities and stakeholders.** After framing the challenge and exploring early ideas, expand the process to include community members, frontline workers, or other key stakeholders. Sessions can be held separately or in mixed groups, depending on the context, and should be designed to make all participants feel comfortable contributing. Practical strategies to encourage participation include:
    - **Starting simple.** Use low-risk icebreakers, such as stickers on a picture to show feelings or a quick sketch of a daily routine. These tasks warm up the group and reduce fear of being "wrong."



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Idea Generation 1A: Brainstorming session

Use this worksheet to guide a brainstorming session with your team or a co-creation session with community members and people with lived experience. Focus on surfacing small behaviours, challenges, and practical ideas through hands-on, visual, and open-ended activities.

Preparation

Steps to complete in preparation for the co-creation session

- Brainstorming: Invite a few colleagues to join. Diversity is good, we want different perspectives.
- Co-creation: Invite 4–8 people from the community
- Plan for 60–90 minutes
- Choose a comfortable and inspiring space

Brainstorming tips

Keep these in mind

- Set a timer: 15–20 min per HMW prompt.
- Bring real examples, props, visual prompts, or stories.
- Use sticky notes and markers. Write one idea per sticky note.
- Don't force ideas, let people reflect and warm up.
- Encourage quantity over quality, don't judge ideas too early. Go for wild, refine later!
- Capture ideas with drawings or direct quotes.
- Build on each other's ideas: use "yes, and..." Treat all input as valuable, even if it's rough or contradictory

How Might We ...

Pull the HMWs developed in the previous activity. During the session, do a different activity for each HMW.

Brainstorming activity

Use interactive Brainstorming techniques to get participants thinking about solutions.

Fix-it Cards

Hand out cards that say: "It would be easier if..." and let people finish the sentence for the prompt.

Draw the idea

Ask people to draw what could make it easier, clearer, or more welcoming. No art skills needed!

Yes, and...

Start with one idea, then go around and add to it with "Yes, and..." to build it out.

Act it out

Have participants act out what currently happens, then "replay" with a small change.

Remix and match

Show flyers, posters, or tools used in the past and ask: "What would you change or combine to make this better?"

Brick-by-Brick

Start with the question: "What's one small thing that could help?" Write each response on a sticky note and stack them like building blocks.

- **Using local language and cultural formats.** Conduct discussions in the language participants are most comfortable with, and lean on familiar modes of expression (role play, storytelling, songs) to make participation natural.
- **Setting ground rules clearly and visually.** Co-create rules (e.g., "one person speaks at a time," "all ideas are valid") and keep them visible to reinforce safety and respect.
- **Actively managing power dynamics.** Design activities that create space for quieter or less powerful voices. For example, begin with a small group or pair discussions before moving into plenary, or rotate individuals presenting ideas so youth or marginalized voices can take the lead. Where appropriate, consider holding separate sessions so participants can speak more openly without pressure from others.
- **Offering multiple ways to contribute.** Provide materials like cards, sticky notes, or objects so participants can write, draw, or place ideas instead of speaking out loud — giving options for those less comfortable speaking in groups.
- **Modelling openness.** Facilitators can share a story or admit not knowing an answer to signal that vulnerability is welcome and honesty is safe.

Co-creation works best when people feel equal, respected, and safe to share their ideas. The role of the facilitator is not to provide answers, but to create the conditions where all participants can collaborate, build on one another's perspectives, and imagine new possibilities. To facilitate a co-creation, follow the guidance below:

a. Choose three to four concrete "how might we" questions from the Ideation Prompts worksheet to guide the session.

- Write each "how might we" question clearly on a separate poster or flipchart sheet. Display these around the room, leaving space for idea generation.
- Give participants a stack of sticky notes and markers.

b. Encourage curiosity, active listening, and collaboration so participants feel free to contribute without judgment. Introduce simple rules like "yes, and…" thinking, where participants build on each other's ideas, rather than dismissing them. Remind the group that no idea is too small or too far-fetched — many of the strongest interventions emerge from combining the familiar with the unexpected.

c. Bring key research insights and aids that can help to spark ideas for interventions and guide the discussion. Visuals like behavioural maps or profiles can be useful as they often reveal where support is most needed, while quotes and field observations can inspire creative solutions based on real challenges and moments.

d. Begin facilitating the idea generation session by focusing on one "how might we" question at a time. Invite participants to share ideas verbally or write them down individually — anything that could help to address the prompt. Capture one idea per sticky note, whether contributed aloud or written directly by participants.

e. Rotate through prompts. After about 10–15 minutes of brainstorming for the first prompt, rotate to the next "how might we" question and repeat the process. Continue until participants have had a chance to contribute intervention ideas under each prompt.

f. To help participants generate ideas more easily — especially those who may feel unsure where to begin — facilitators can introduce the following activities, designed to unlock creativity and fresh thinking around the "how might we" questions. These can be rotated across prompts, or set up as stations around the room for participants to explore.

TABLE 2. BRAINSTORMING ACTIVITIES

ACTIVITY	DESCRIPTION	USEFUL FOR	INSTRUCTIONS
Fix-it Cards	Participants complete the sentence "The behaviour would be easier if..." to identify simple, practical changes that could remove barriers or make a behaviour more doable. This activity helps to surface low-effort, high-impact ideas rooted in participants' real experiences.	Breaking down barriers into actionable solutions, especially for those less familiar with formal brainstorming.	1. Distribute cards or sticky notes. 2. Ask participants to finish the sentence: "The behaviour would be easier if..." 3. Share and group similar suggestions.
Draw the Idea	Participants sketch simple visuals to communicate their ideas. No art skills needed.	Engaging participants who think better visually, or have lower literacy levels.	1. Hand out paper or sticky notes. 2. Ask: "What could help make this easier?" 3. Invite quick sketches. 4. Ask them to describe their idea in their own words.
Yes, and...	Participants build on ideas collaboratively, with everyone encouraged to add.	Generating energy and group ownership, or expanding initial thoughts into richer concepts.	1. Start with one idea. 2. Each person adds with "Yes, and..." 3. Build a chain of suggestions. 4. Capture the evolving idea.
Act it Out	Participants role-play what happens now, and then what could happen with a small change.	Testing realism and emotional response, or engaging kinesthetic learners.	1. Choose a real-life scenario. 2. Role-play the current experience. 3. Re-play it with a simple improvement. 4. Reflect as a group.
Remix and Match	Participants revise real-world examples of past materials.	Stimulating redesign thinking, especially for caregivers or workers familiar with local materials.	1. Show posters, flyers, or tools. 2. Ask: "What would you change or combine?" 3. Encourage tweaks or mashups. 4. Sketch or describe the new version.
Brick-by-Brick	Participants contribute small ideas that build into a bigger solution.	Helping groups that struggle with big-picture thinking, along with promoting inclusiveness.	1. Ask: "What's one small thing that could help?" 2. Write one idea per sticky note. 3. Stack or arrange them visually. 4. Review and build a concept.

g. Each "how might we" poster will now have a cluster of sticky notes beneath it, creating a visual, participatory pool of early intervention concepts. Document the outputs in a spreadsheet or document. For every idea, note:

- What the idea is
- Which "how might we" question the idea responds to
- What behavioural barrier, enabler, or insight it links to, or barrier the idea addresses
- The community's perceived value of the idea
- Any supporting quotes or observations that illustrate why the idea matters

For example, if an idea is "send SMS reminders," this could be supported by a participant's quote, such as "I often forget appointments unless someone reminds me the day before." Or, if the idea is "move registration to the marketplace," this could connect to an observation that mothers consistently mentioned passing through the market every day, while rarely going to the clinic. Capturing this context preserves the reasoning behind ideas, making it easier to spot patterns or clusters when refining them later.

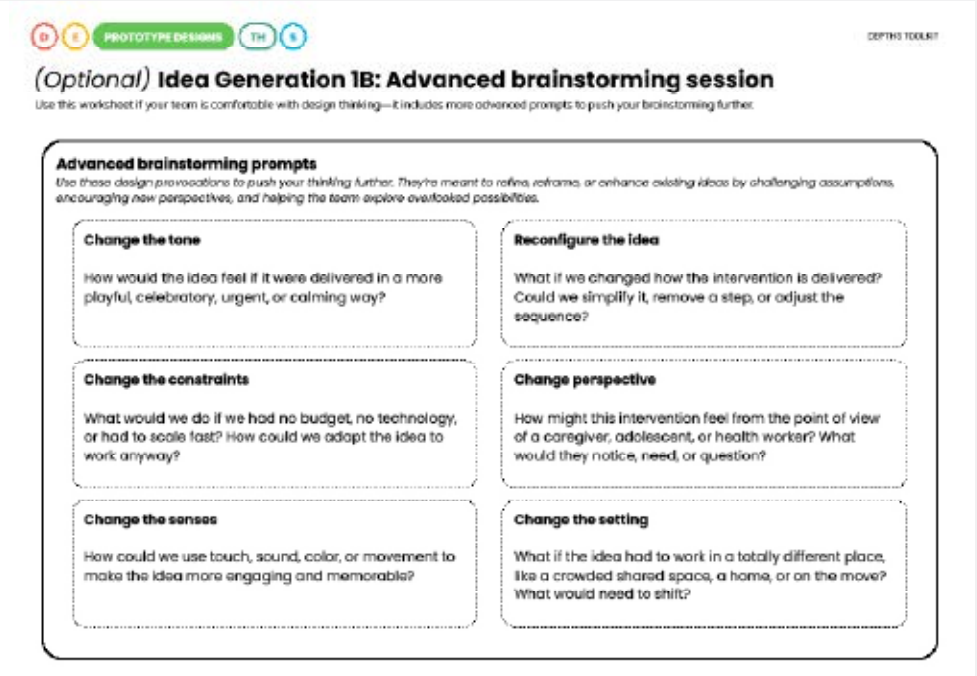
2.2. Optional brainstorming activity to refine ideas: Design Provocation

As an optional extension of the co-creation session, the **Design Provocations** technique helps strengthen ideas by encouraging teams to rethink how an intervention could work or be experienced. It is especially useful when early ideas feel too obvious or incomplete.

Design provocations use "what if" questions to revisit existing ideas with fresh eyes. The goal isn't to replace earlier concepts, but to deepen, adapt, or transform them into stronger, more innovative, and context-appropriate solutions.

a. Start by reviewing the table created at the end of the co-creation session, which links each idea to its original "how might we" question.

b. Then, choose provocation prompts and apply them to each idea to see how it might evolve. Table 3 provides a list of provocation prompts, alongside concrete examples to show how ideas can evolve into more relevant designs.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

(Optional) **Idea Generation 1B: Advanced brainstorming session**

Use this worksheet if your team is comfortable with design thinking—it includes more advanced prompts to push your brainstorming further.

Advanced brainstorming prompts

Use these design provocations to push your thinking further. They're meant to refine, reframe, or enhance existing ideas by challenging assumptions, encouraging new perspectives, and helping the team explore overlooked possibilities.

Change the tone

How would the idea feel if it were delivered in a more playful, celebratory, urgent, or calming way?

Reconfigure the idea

What if we changed how the intervention is delivered? Could we simplify it, remove a step, or adjust the sequence?

Change the constraints

What would we do if we had no budget, no technology, or had to scale fast? How could we adapt the idea to work anyway?

Change perspective

How might this intervention feel from the point of view of a caregiver, adolescent, or health worker? What would they notice, need, or question?

Change the senses

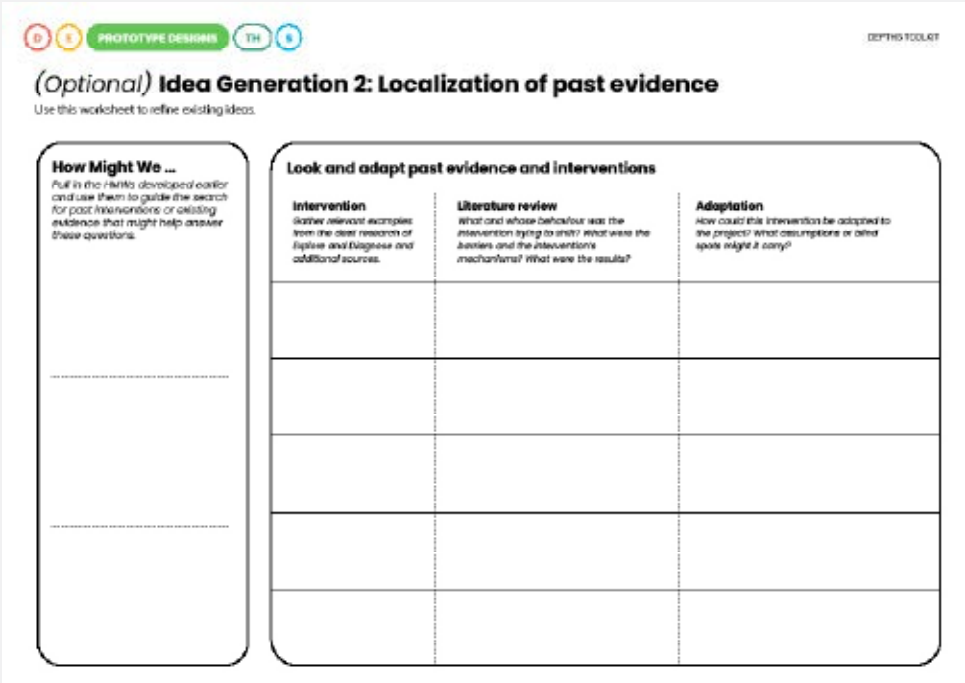
How could we use touch, sound, color, or movement to make the idea more engaging and memorable?

Change the setting

What if the idea had to work in a totally different place, like a crowded shared space, a home, or on the move? What would need to shift?

TABLE 3. DESIGN PROVOCATION PROMPTS

PROVOCATION	PROMPT	USEFUL FOR	EXAMPLE: ORIGINAL IDEA	EXAMPLE: REDESIGNED IDEA
Change the tone	What if the idea were more playful, celebratory, urgent, or reassuring?	Making the idea more emotionally resonant or better aligned with audience values	A plain reminder card with vaccine appointment date	A cheerful "Vaccine Hero" card with celebratory language and a sticker for the child
Reconfigure the idea	What if we delivered the intervention differently? Could we change the order, remove a step, or simplify it?	Simplifying the delivery or reducing logistical barriers	Text reminder sent two days before appointment	A physical reminder given at the last appointment, attached to the child's health booklet for ongoing visibility
Change the constraints	What if we had no money, no tech, or had to scale quickly? How would we adapt?	Adapting ideas to low-resource or rapidly scaling environments	A digital appointment reminder system	Health workers use community radio announcements to remind caregivers of upcoming vaccine days
Change perspective	What if a caregiver experienced this intervention? What would we need to consider? What about an adolescent? A health care worker?	Ensuring ideas are user-centred and account for varying needs and experiences	A group education session on vaccine schedules	Peer mothers share their own stories about vaccination journeys in small, informal conversations
Change the senses	What if we engaged touch, sound, colour, or movement to make the idea more memorable?	Making ideas more engaging and accessible, especially for children or low-literacy populations	A printed immunization leaflet handed out at the clinic	A colourful, illustrated wall poster with movable vaccine dose markers placed in waiting areas



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

(Optional) **Idea Generation 2: Localization of past evidence**

Use this worksheet to refine existing ideas.

How Might We ...

Pull in the HMWs developed earlier and use them to guide the search for past interventions or existing evidence that might help answer these questions.

Look and adapt past evidence and interventions

Intervention *Gather relevant examples from the desk research of Explore and Diagnose and additional sources.*	**Literature review** *What and whose behaviour was the intervention trying to shift? What were the barriers and the intervention's mechanisms? What were the results?*	**Adaptation** *How could this intervention be adapted to the project? What assumptions or blind spots might it carry?*

2.3. Optional idea generation activity: Localization of past evidence

Strong ideas don't always need to start from scratch. In many cases, similar behavioural challenges have been tackled elsewhere, and these existing programmes, studies, or interventions can offer inspiration. This activity reviews evidence to see whether proven approaches can be adapted to the current challenge. This might include:

- Interventions documented in research studies or evaluations
- Programmes implemented in similar regions or populations
- Case studies from global evidence platforms or organizational archives
- Insights shared by implementing partners or technical advisors

The goal is not to copy, but to identify core behavioural strategies that worked in other settings, then adapt them for the current local context. Consider: What made the intervention effective elsewhere? What would need to change for it to work here? This is especially useful under time or resource constraints, or when building on past successes.

a. Begin by revisiting the "how might we" questions and search for interventions and past evidence that could help answer them.

b. Use findings from the initial desk research conducted in the Explore & Diagnose phase to identify past interventions related to the problem and the behavioural barriers. Review this evidence, and consider:

- What and whose behaviour was the intervention trying to shift?
- What were the barriers and the intervention's mechanisms?
- What were the results?

c. Reflect on localization and adaptation by asking:

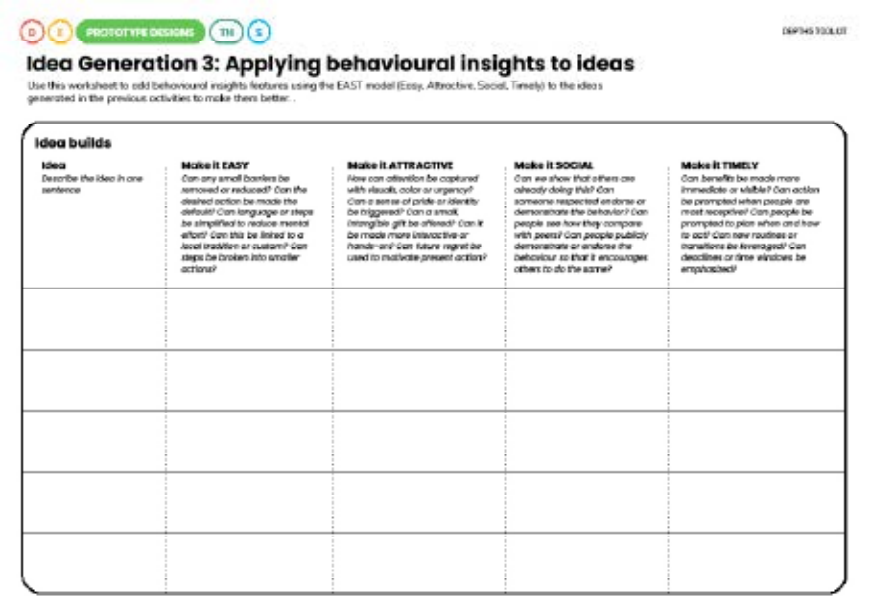
- How could this intervention be adapted to the project?
- What assumptions or blind spots might it carry?

While leveraging and using past evidence is an efficient and effective approach, it is also important to explore new ideas.

2.4. Apply a behavioural science framework to refine intervention ideas: EAST framework

This step helps ensure intervention ideas are both creative and practical. Applying the EAST framework is a required step in the DEPTHS process, as it allows ideas to be assessed and refined using principles of behavioural science. To balance creativity with rigour, it is recommended to complete at least two brainstorming exercises before applying EAST.

While several behavioural science frameworks exist—such as MINDSPACE or the Behaviour Change Wheel—this field guide uses the EAST framework for its practicality and wide adoption.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Idea Generation 3: Applying behavioural insights to ideas

Use this worksheet to add behavioural insights features using the EAST model (Easy, Attractive, Social, Timely) to the ideas generated in the previous activities to make them better.

Idea builds

Idea *Describe the idea in one sentence*	Make it EASY *Can any small barriers be removed or reduced? Can the desired action be made the default? Can language or steps be simplified to reduce mental effort? Can this be linked to a local tradition or custom? Can steps be broken into smaller actions?*	Make it ATTRACTIVE *How can attention be captured with visuals, color or urgency? Can a sense of pride or identity be triggered? Can a small, intangible gift be offered? Can it be made more interactive or hands-on? Can future regret be used to motivate present action?*	Make it SOCIAL *Can we show that others are already doing this? Can someone respected endorse or demonstrate the behavior? Can people see how they compare with peers? Can people publicly demonstrate or endorse the behaviour so that it encourages others to do the same?*	Make it TIMELY *Can benefits be made more immediate or visible? Can action be prompted when people are most receptive? Can people be prompted to plan when and how to act? Can new routines or transitions be leveraged? Can deadlines or time windows be emphasized?*

E	A	S	T
Easy	Attractive	Social	Timely

BEHAVIOURAL DESIGN SPOTLIGHT:

The EAST Framework

About the Framework

The EAST framework was developed by the Behavioural Insights Team (BIT)—also known as the "Nudge Unit"—as a practical, accessible way to apply behavioural science in public policy and program design. It synthesizes decades of behavioural economics and cognitive psychology research into four core principles: Easy, Attractive, Social, and Timely. Each principle corresponds to a well-documented behavioural barrier or opportunity, helping designers craft interventions that are both human-centred and evidence-informed. EAST is flexible and especially useful during ideation, prototyping, or when turning insights into strategies.

The Four Principles, Explained

EASY

Barrier addressed: People often avoid tasks that feel complicated, time-consuming, or mentally taxing.

Behavioural insight: Humans are cognitive misers: our brains default to the easiest available option.

What to do:

- Simplify steps, reduce complexity, and eliminate unnecessary choices
- Use plain language and intuitive visuals
- Pre-fill forms or automate processes when possible
- Make the desired behaviour the default

Example: In cash transfer programs, shifting from paper vouchers to mobile money made it significantly easier for families to access support, eliminating long travel times, reducing paperwork, and allowing funds to be received and used with just a few clicks. A randomized experiment in Niger found that households receiving mobile transfers had higher dietary diversity and children consumed more meals per day, benefits largely attributed to time saved from traveling and waiting, as well as enhanced decision-making power for women.[2]

ATTRACTIVE

Barrier addressed: Competing stimuli and low motivation make it hard to capture attention.

Behavioural insight: People are more likely to engage with environments that stand out or feel rewarding.

What to do:

- Use colour, imagery, and visual cues to make options stand out
- Frame messages in ways that resonate emotionally or convey personal benefit
- Introduce small incentives or commitments that spark action

Example: In handwashing campaigns, painting bright footprints on the ground that lead children from the latrine to the handwashing station turned the behaviour into a game-like experience. The colourful visual cues caught children's attention and made handwashing feel fun and rewarding, increasing consistent use[3].

SOCIAL

Barrier addressed: People are influenced by what others around them do or believe.

Behavioural insight: Social norms, peer behaviour, and group identity strongly shape decisions.

What to do:

- Make positive behaviours visible and relatable
- Highlight stories from peers or trusted community members
- Use group pledges, social accountability, or peer reminders

Example: In girls' education programs in India, showing parents stories and posters of local female role models who had successfully completed school shifted perceptions about the value of girls' education. Parents were more likely to see supporting their daughters' schooling as both common and celebrated within their own community[4].

2 Aker, J.C., Boumnijel, R., McClelland, A. and Tierney, N., 2014. Zap it to me: The short-term impacts of a mobile cash transfer program. Discussion Paper No. 268. Paris: Agence Française de Développement (AFD). Available at: https://www.calpnetwork.org/wp-content/uploads/2020/01/zap-26aug2014-1.pdf [Accessed 2 September 2025].

3 Dreibelbis, R., Kroeger, A., Hossain, K., Venkatesh, M. and Ram, P.K., 2016. Behavior Change without Behavior Change Communication: Nudging Handwashing among Primary School Students in Bangladesh. International Journal of Environmental Research and Public Health, 13(1), p.129

4 Beaman, L., Duflo, E., Pande, R. and Topalova, P., 2012. Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. Science, 335(6068), pp.582–586.

TIMELY

Barrier addressed: Even with good intentions, people often delay or forget to act.

Behavioural insight: Timing influences readiness and follow-through.

What to do:

- Time interventions around key decision moments or life events
- Send reminders close to the moment of action
- Use planning tools to prompt commitment (e.g., calendars, SMS reminders)

Example: In Sudan, integrating birth registration services directly within health centres made it easy for parents to register their newborns at the time of delivery. By aligning the service with the moment parents were already completing health paperwork, registration became immediate and convenient. As a result, more than 170,500 newborns were registered in remote localities in a single year.[5]

Why EAST Works

EAST works because it aligns with how people actually make decisions, which is often emotional, habitual, and context-driven, rather than rational or linear. Many behavioural interventions fail because they assume knowledge or awareness automatically leads to action. EAST helps to counter this by designing for the realities of human behaviour, using small nudges to steer choices without coercion.

It's not a checklist, but a lens. Not every intervention will need all four principles, but applying at least one often makes an idea stronger.

5 UNICEF, 2016. Annual Results Report 2016: Health. New York: UNICEF. Available at: https://www.unicef.org/media/49126/file/2016arr_health.pdf [Accessed 2 September 2025].

a. Review the spreadsheet or table where ideas from the brainstorming session were documented.

b. Transfer selected ideas. If there are many, select the most promising ideas and list them (one per row) in the EAST worksheet.

c. Use the EAST prompts as creative triggers to explore how each behavioural principle might improve the idea. For example:

- **Easy:** Can steps be simplified or linked to familiar routines?
- **Attractive:** Could visuals, rewards, or emotional appeals help?
- **Social:** Is there a way to show that others are doing it too?
- **Timely:** Are we catching people at the right moment?

d. In each EAST column of the worksheet (Make it Easy, Make it Attractive, Make it Social, and Make it Timely) write down refinements, enhancements, or creative twists based on the prompts.

e. Where possible, combine principles and levers: strong ideas often touch more than one. Table 5 presents further guidance on how to apply the EAST framework to improve ideas.

TABLE 4. PROMPTS FOR IDEA GENERATION FOLLOWING THE EAST FRAMEWORK

CATEGORY	BEHAVIOURAL MECHANISM	PROMPT	ORIGINAL IDEA	NEW IDEA EXAMPLE
EASY	Friction costs	Can any small barriers (e.g., travel time, forms, wait) be removed or reduced?	Hold weekly vaccination days at health clinics.	Offer mobile vaccination clinics near marketplaces to reduce travel time and increase convenience for caregivers.
	Defaults	Can the desired action be made the default?	Ask parents during health visits if they want to opt into vaccine reminders.	Automatically enrol children in routine immunization reminders upon birth registration, unless parents opt out.
	Cognitive load reduction	Can language or steps be simplified to reduce mental effort?	Provide a printed list of upcoming vaccine dates.	Use illustrated vaccine cards with icons instead of text for caregivers with low literacy.
	Cultural alignment	Can this be linked to a local tradition or custom?	Plan vaccination events during school holidays.	Schedule vaccination drives during local festivals where families already gather, aligning with community rhythms.
	Chunking information	Can steps be broken into smaller actions?	Distribute a pamphlet with all vaccine information.	Give caregivers a one-step instruction card for 'what to do before' and 'what to expect after' a vaccine visit.
ATTRACTIVE	Salience	How can attention be captured with visuals, colour or urgency?	Send standard text message reminders about appointments.	Design bright, eye-catching appointment slips for mothers with a visual countdown to the next vaccine date.
	Emotional rewards	Can a sense of pride or identity be triggered?	Give verbal thanks to caregivers after a vaccination.	Offer caregivers a 'Health Hero' badge for completing all childhood vaccines, reinforcing parental pride.
	Reciprocity	Can a small, intangible gift be offered?	Say "thank you" after each vaccine appointment.	After each vaccine visit, provide a sticker or thank-you card recognizing caregivers' commitment to child health.
	Experiential engagement	Can it be made more interactive or hands-on?	Place posters about vaccines in the waiting area.	Set up a vaccination play corner for children to reduce fear and make the experience engaging.
	Anticipated regret	Can future regret be used to motivate present action?	Inform caregivers of the vaccine's importance.	Use messages like "You may regret missing today's vaccine—protect your child now" to evoke anticipated emotion.

Table continues on following page.

TABLE 4 [CONTINUED]. PROMPTS FOR IDEA GENERATION FOLLOWING THE EAST FRAMEWORK

CATEGORY	BEHAVIOURAL MECHANISM	PROMPT	ORIGINAL IDEA	NEW IDEA EXAMPLE
SOCIAL	Social norms	Can we show that others are already doing this?	Encourage parents to vaccinate through announcements.	Share data that "85% of mothers in your community have vaccinated their children this year."
	Messenger effect	Can someone respected endorse or demonstrate the behaviour?	Ask parents to encourage each other to vaccinate.	Have respected religious or community leaders vaccinate their children publicly and share the moment.
	Peer comparison	Can people see how they compare with peers?	Remind parents about routine immunization schedules.	Send caregivers a card noting how many other families in their neighbourhood have completed vaccines.
	Social proof	Can people publicly demonstrate or endorse the behaviour so that it encourages others to do the same?	Congratulate parents quietly after vaccination.	Give colourful bracelets every time a caregiver vaccinates their child so they can signal their participation to others.
TIMELY	Present bias	Can benefits be made more immediate or visible?	Explain that vaccines prevent future illness.	Highlight the immediate relief of protection post-vaccine (e.g., "Your child is protected today, not just later").
	Timing alignment	Can action be prompted when people are most receptive?	Send reminders a week before appointments.	Send SMS reminders in the early morning when caregivers typically plan their day.
	Planning prompts	Can people be prompted to plan when and how to act?	Give a paper with vaccine schedules.	Include a blank calendar space on vaccine cards for caregivers to write their own reminder date and time.
	Moments of change	Can new routines or transitions be leveraged?	Invite families to vaccinate at any time.	Bundle vaccine registration with school enrolment, when families are already thinking about child health.
	Time scarcity	Can deadlines or time windows be emphasized?	Share dates for vaccine availability.	Use urgent messages like "Only 3 days left for the free measles vaccine clinic in your area!"

3. Consolidate intervention ideas

a. After brainstorming and refinement, compile ideas from the **EAST** worksheet into one list. Remove duplicates and set aside any that:

- Don't target the priority behaviour
- Fall outside the scope or objectives of the project
- Would require entirely new systems, or actors not included in the design process

For example: An idea like "Engaging agricultural extension workers to promote vaccination during home visits" may have merit, but it might fall outside the scope if the current intervention is focused on urban clinic settings, with no existing relationship to the agriculture sector.

b. Next, merge ideas that are similar or complementary. Some may share the same behavioural mechanism, delivery channel, or target group and can be strengthened when brought together. For example:

- **Idea 1:** Send reminder messages signed by a health worker
- **Idea 2:** Inform caregivers that a vaccine has already been reserved for them
- **Merged idea:** Send reminder messages signed by a health worker, informing caregivers that a vaccine has been reserved in their name.

c. The final output of this step should be a curated list of distinct, behavioural science informed, grounded ideas that:

- Address the prioritized behaviour
- Fit the local context
- Apply behavioural insights
- Are feasible within project scope

This consolidated list will become the foundation for prioritization and prototyping in the next step.

Consider the following tips when generating ideas:

- **Keep "just good enough" versions of ideas.** While they may be easy to ignore, some of the best solutions evolve from rough sketches. Instead of filtering these ideas out too early, capture them in a "maybe" pile — a separate column or sheet to revisit them later. When refined, they often address gaps others missed.
- **Map ideas to the micro-behaviour timeline.** After brainstorming, take the top 10–15 ideas and map each one to the behavioural timeline (e.g., **before**, **during**, or **after** the target action). This helps to check whether your solutions are too clustered around one moment (e.g., reminders before the appointment) and identifies neglected stages where ideas are still needed.
- **Run a quick "would this be useful to me?" test.** For each intervention idea, ask someone from the target population — or role-play as them — and answer: "Would this actually help me do the behaviour?" This gut-check helps to surface ideas that sound good on paper, but don't feel helpful in practice. If the answer is "not really," refine or rethink.

CASE STUDY:

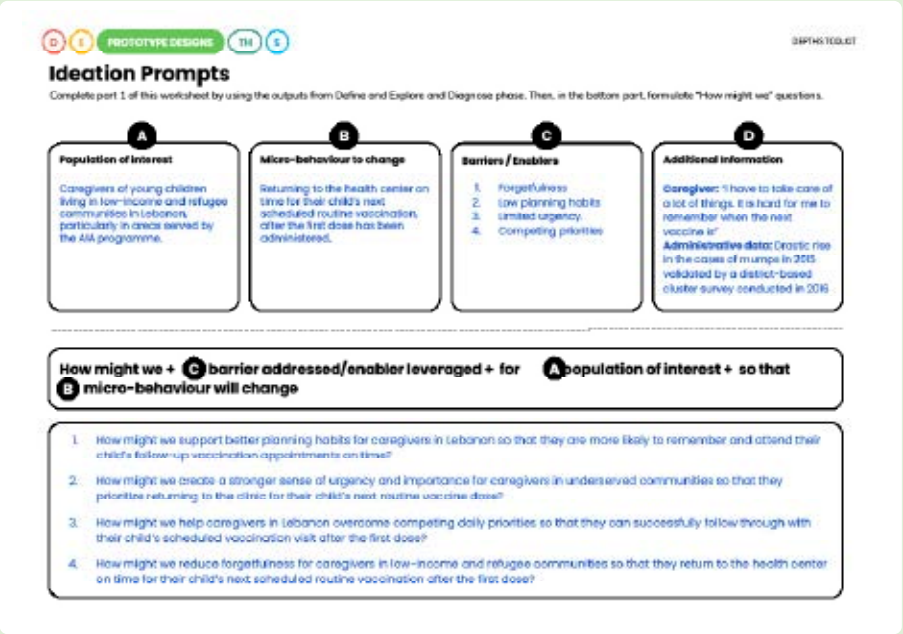
Increasing childhood vaccination uptake in Lebanon

These Ideation Prompts and Idea Generation worksheets were not developed by the original project team. They are recreated examples based on real project data and context.

In Lebanon, the project team set out to increase childhood vaccination uptake among refugee caregivers served by the Ministry of Public Health's AIA programme. Drawing on insights from the **Explore & Diagnose** phase, the team first synthesized key findings into a structured "how might we" worksheet. This helped to distill complex research into focused prompts, by specifying the target micro-behaviour and identifying barriers. One of the most important prompts to emerge was: "How might we reduce forgetfulness and lack of planning for refugee caregivers so that they return to the health centre on time for their child's next scheduled vaccine?"

This question served as the foundation for a number of intervention ideas generated through co-creation and team ideation. Some initial concepts included:

- A simple paper appointment card that health workers would hand to caregivers after the child's first vaccine.
- A version of that card redesigned to include a space for caregivers to write down the return date, to reinforce commitment and memory.
- A reminder message signed by a known health worker, sent a few days before the next appointment.
- A visual calendar sticker caregivers could use at home to track the date.
- A verbal commitment moment, where caregivers would state aloud when they planned to return.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Ideation Prompts

Complete part 1 of this worksheet by using the outputs from Define and Explore and Diagnose phase. Then, in the bottom part, formulate "How might we" questions.

A — Population of interest

Caregivers of young children living in low-income and refugee communities in Lebanon, particularly in areas served by the AIA programme.

B — Micro-behaviour to change

Returning to the health center on time for their child's next scheduled routine vaccination, after the first dose has been administered.

C — Barriers / Enablers

1. Forgetfulness
2. Low planning habits
3. Limited urgency
4. Competing priorities

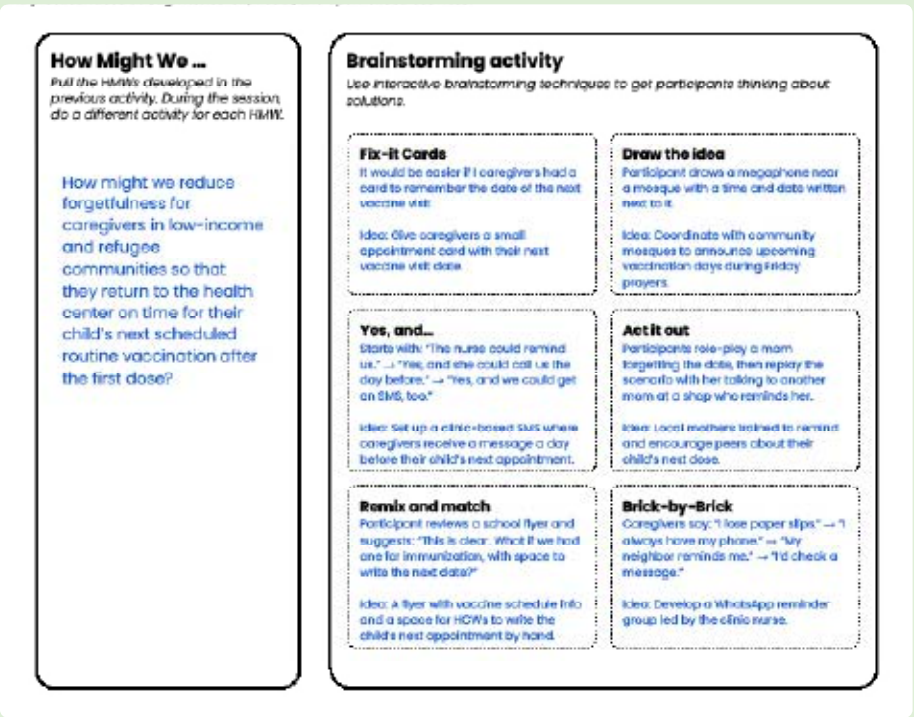
D — Additional information

Caregiver: "I have to take care of a lot of things. It is hard for me to remember when the next vaccine is"

Administrative data: Drastic rise in the cases of mumps in 2015 validated by a district-based cluster survey conducted in 2016

How might we + C barrier addressed/enabler leveraged + for A population of interest + so that B micro-behaviour will change

1. How might we support better planning habits for caregivers in Lebanon so that they are more likely to remember and attend their child's follow-up vaccination appointments on time?
2. How might we create a stronger sense of urgency and importance for caregivers in underserved communities so that they prioritize returning to the clinic for their child's next routine vaccine dose?
3. How might we help caregivers in Lebanon overcome competing daily priorities so that they can successfully follow through with their child's scheduled vaccination visit after the first dose?
4. How might we reduce forgetfulness for caregivers in low-income and refugee communities so that they return to the health center on time for their child's next scheduled routine vaccination after the first dose?



How Might We ...

Pull the HMWs developed in the previous activity. During the session, do a different activity for each HMW.

How might we reduce forgetfulness for caregivers in low-income and refugee communities so that they return to the health center on time for their child's next scheduled routine vaccination after the first dose?

Brainstorming activity

Use interactive brainstorming techniques to get participants thinking about solutions.

Fix-it Cards

It would be easier if I caregivers had a card to remember the date of the next vaccine visit

Idea: Give caregivers a small appointment card with their next vaccine visit date

Draw the idea

Participant draws a megaphone near a mosque with a time and date written next to it

Idea: Coordinate with community mosques to announce upcoming vaccination days during Friday prayers

Yes, and...

Starts with: "The nurse could remind us." → "Yes, and she could call us the day before." → "Yes, and we could get an SMS, too."

Idea: Set up a clinic-based SMS where caregivers receive a message a day before their child's next appointment.

Act it out

Participants role-play a mom forgetting the date, then replay the scenario with her talking to another mom at a shop who reminds her.

Idea: Local mothers trained to remind and encourage peers about their child's next dose.

Remix and match

Participant reviews a school flyer and suggests: "This is clear. What if we had one for immunization, with space to write the next date?"

Idea: A flyer with vaccine schedule info and a space for HCWs to write the child's next appointment by hand.

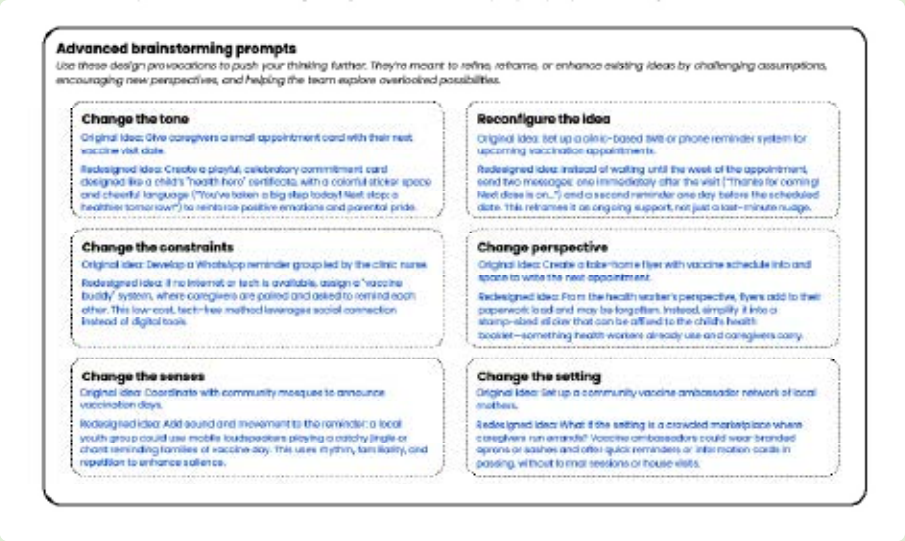
Brick-by-Brick

Caregivers say: "I lose paper slips." → "I always have my phone." → "My neighbor reminds me." → "I'd check a message."

Idea: Develop a WhatsApp reminder group led by the clinic nurse.

These early-stage ideas were generated using tools such as Fix-it Cards, Act It Out, and Brick-by-Brick, enabling the team to collaboratively explore small yet impactful design changes. The appointment card — later refined and tested — emerged as the most promising solution.

To strengthen these ideas, the team also reviewed past evidence showing the effectiveness of SMS reminders and physical cards in boosting vaccine uptake. This encouraged a focus on low-tech solutions that were affordable, feasible, and aligned with caregivers' preferences and context.



Advanced brainstorming prompts

Use these design provocations to push your thinking further. They're meant to refine, reframe, or enhance existing ideas by challenging assumptions, encouraging new perspectives, and helping the team explore overlooked possibilities.

Change the tone

Original idea: Give caregivers a small appointment card with their next vaccine visit date.

Redesigned idea: Create a playful, celebratory commitment card designed like a child's 'health hero' certificate, with a colorful sticker space and cheerful language ("You've taken a big step today! Next stop: a healthier tomorrow") to reinforce positive emotions and parental pride.

Reconfigure the idea

Original idea: Set up a clinic-based SMS or phone reminder system for upcoming vaccination appointments.

Redesigned idea: Instead of waiting until the week of the appointment, send two messages: one immediately after the visit ("Thanks for coming! Next dose is on...") and a second reminder one day before the scheduled date. This reframes it as ongoing support, not just a last-minute nudge.

Change the constraints

Original idea: Develop a WhatsApp reminder group led by the clinic nurse.

Redesigned idea: If no internet or tech is available, assign a "vaccine buddy" system, where caregivers are paired and asked to remind each other. This low-cost, tech-free method leverages social connection instead of digital tools.

Change perspective

Original idea: Create a take-home flyer with vaccine schedule info and space to write the next appointment.

Redesigned idea: From the health worker's perspective, flyers add to their paperwork load and may be forgotten. Instead, simplify it into a stamp-sized sticker that can be affixed to the child's health booklet—something health workers already use and caregivers carry.

Change the senses

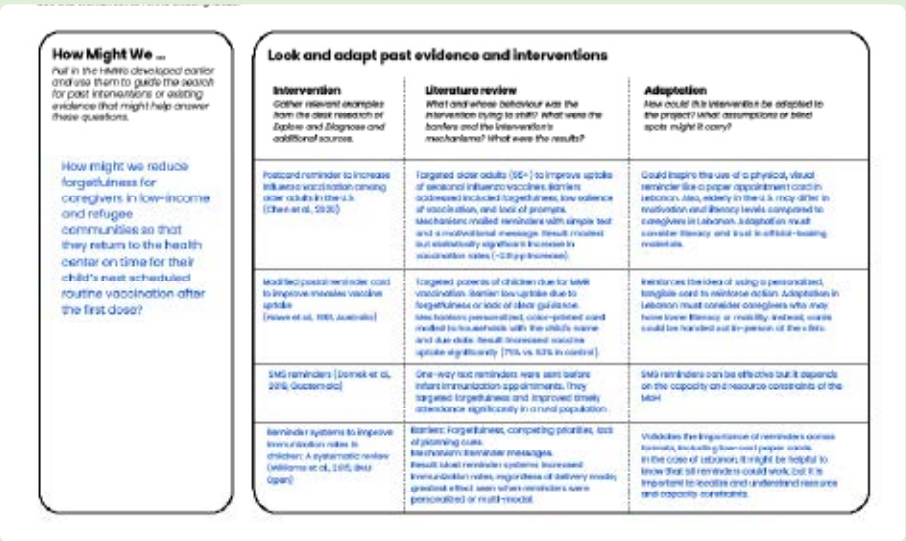
Original idea: Coordinate with community mosques to announce vaccination days.

Redesigned idea: Add sound and movement to the reminder: a local youth group could use mobile loudspeakers playing a catchy jingle or chant reminding families of vaccine day. This uses rhythm, familiarity, and repetition to enhance salience.

Change the setting

Original idea: Set up a community vaccine ambassador network of local mothers.

Redesigned idea: What if the setting is a crowded marketplace where caregivers run errands? Vaccine ambassadors could wear branded aprons or washes and offer quick reminders or information cards in passing, without formal sessions or house visits.



How Might We ...

Pull in the insights developed earlier and use them to guide the search for past interventions or existing evidence that might help answer these questions.

How might we reduce forgetfulness for caregivers in low-income and refugee communities so that they return to the health center on time for their child's next scheduled routine vaccination after the first dose?

Look and adapt past evidence and interventions

Intervention *Gather relevant examples from the desk research or Explore and Diagnose and additional sources.*	Literature review *What and where behaviour was the intervention trying to shift? What were the barriers and the intervention's mechanisms? What were the results?*	Adaptation *How could this intervention be adapted to the project? What assumptions or blind spots might it carry?*
Postcard reminder to increase influenza vaccination among older adults in the U.S. (Chen et al., 2020)	Targeted older adults (65+) to improve uptake of seasonal influenza vaccines. Barriers addressed included forgetfulness, low salience of vaccination, and lack of prompts. Mechanisms: mailed reminders with simple text and a motivational message. Result: modest but statistically significant increase in vaccination rates (~2.8 pp increase).	Could inspire the use of a physical, visual reminder like a paper appointment card in Lebanon. Also, elderly in the U.S. may differ in motivation and literacy levels compared to caregivers in Lebanon. Adaptation must consider literacy and trust in official-looking materials.
Modified postal reminder card to improve measles vaccine uptake (Mawe et al., 1999, Australia)	Targeted parents of children due for MMR vaccination. Barrier: low uptake due to forgetfulness or lack of clear guidance. Mechanism: personalized, color-printed card mailed to households with the child's name and due date. Result: increased vaccine uptake significantly (71% vs. 63% in control).	Reinforces the idea of using a personalized, tangible card to reinforce action. Adaptations in Lebanon must consider caregivers who may have lower literacy or mobility. Instead, cards could be handed out in-person at the clinic.
SMS reminders (Domek et al., 2016, Guatemala)	One-way text reminders were sent before infant immunization appointments. They targeted forgetfulness and improved timely attendance significantly in a rural population.	SMS reminders can be effective but it depends on the capacity and resource constraints of the MOH.
Reminder systems to improve immunization rates in children: A systematic review (Williams et al., 2015, BMJ Open)	Barriers: Forgetfulness, competing priorities, lack of planning cues. Mechanism: Reminder messages. Result: Most reminder systems increased immunization rates, regardless of delivery mode; greatest effect seen when reminders were personalized or multi-modal.	Validates the importance of reminders across formats, including low-cost paper cards. In the case of Lebanon, it might be helpful to know that all reminders could work, but it is important to localize and understand resource and capacity constraints.

The team then applied design provocations and the EAST framework to improve the idea: it was made easier to use by simplifying the language and including visual icons; more attractive through colour and layout; socially reinforced by encouraging caregivers to show it to others; and timed to coincide with critical decision points (e.g. shortly after the initial vaccine).

STEP 2:
Prioritize interventions

In this step:

This step introduces a set of criteria to identify the most promising intervention ideas. Although it's common to recognize the importance of factors such as cost-effectiveness, scalability, and equity, these can be difficult to address during brainstorming sessions. This step provides a structured moment to reflect and make informed decisions. By the end, teams should have a shortlist of 1–3 intervention ideas ready for prototyping and user testing.

Associated tools:

- Idea Prioritization
- Ethics Checklist

Why it matters:

Teams often gravitate toward ideas that feel exciting or familiar, even when they are not feasible or aligned to real barriers. Transparent prioritization criteria encourage balanced assessment and prevent investing in costly, hard-to-scale, or misaligned ideas.

Ethics should be a key part of prioritization. At this stage, the focus is on equity and inclusion in design choices to assess whether intervention ideas risk reinforcing stigma, limiting choice, or excluding vulnerable groups. This kind of "design ethics" check ensures that solutions do no harm and actively promote fairness, dignity, and inclusion.

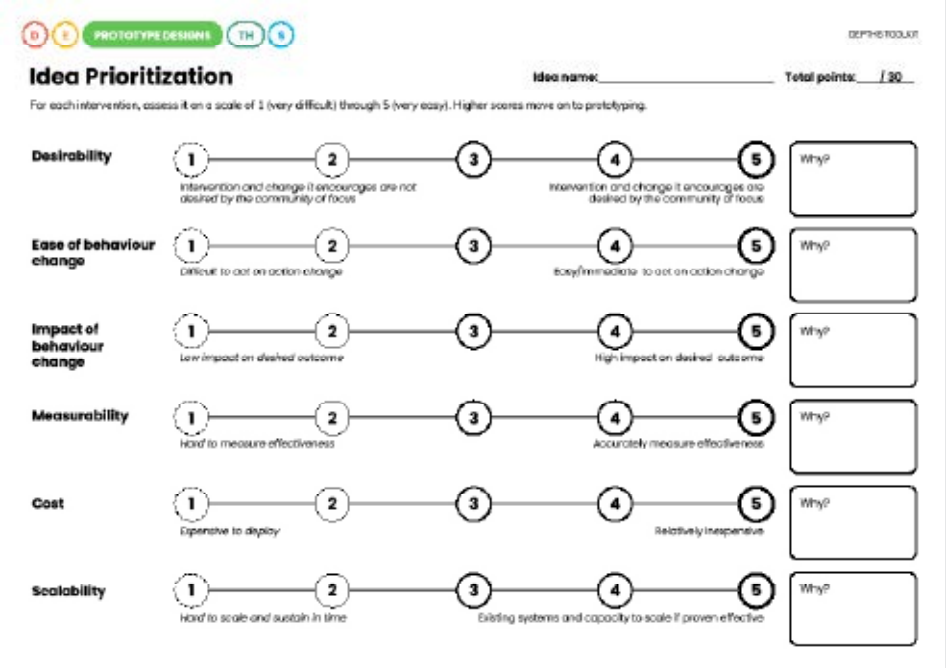
When the process shifts from prioritization to testing ideas in real life, research ethics are introduced to add a further layer. This includes IRB approval, informed consent, safeguarding, and privacy protections — later addressed in detail in the following **Test Hypothesis** chapter.

In this phase, the focus is on prioritizing ideas that are ethically sound in their intent and potential effects; in the next chapter of the DEPTHS process, we ensure that testing those ideas is conducted ethically and responsibly.

How to do it:

1. Prioritize ideas using key criteria

The Idea Prioritization worksheet walks teams through evaluating each idea against a core set of practical criteria. This includes:


D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Idea Prioritization

Idea name: _____ Total points: ___ / 30

For each intervention, assess it on a scale of 1 (very difficult) through 5 (very easy). Higher scores move on to prototyping.

	1	2	3	4	5	Why?
Desirability	Intervention and change it encourages are not desired by the community of focus				Intervention and change it encourages are desired by the community of focus	
Ease of behaviour change	Difficult to act on action change				Easy/immediate to act on action change	
Impact of behaviour change	Low impact on desired outcome				High impact on desired outcome	
Measurability	Hard to measure effectiveness				Accurately measure effectiveness	
Cost	Expensive to deploy				Relatively inexpensive	
Scalability	Hard to scale and sustain in time				Existing systems and capacity to scale if proven effective	

- **Desirability:** Will the idea be accepted and valued by the community?
- **Ease of behaviour change:** Is the idea likely to effectively influence the target behaviour?
- **Impact of behaviour change:** Will a change in this behaviour significantly affect the desired outcome?
- **Measurability:** Can the impact of the idea be measured reliably? Are there credible ways to track effectiveness?
- **Cost:** How resource-intensive is the idea to implement?
- **Scalability:** If successful, can the idea be expanded or adapted to other settings?

a. Create a copy of the **Idea Prioritization** worksheet for each intervention idea from the final list. Write down the name of the idea at the top of the worksheet and assess it across each of the six prioritization criteria listed on the left-hand side. Each criterion should be scored on a scale from 1 (very low) to 5 (very high). The higher the idea's total score, the more suitable the idea is for further development. Table 5 outlines prompts, considerations, and guidance for scoring each criterion, using the reminder appointment card as the example intervention.

TABLE 5: CONSIDERATIONS FOR CRITERIA SCORING

CRITERIA	PROMPT	GUIDE	EXAMPLE
Desirability	Does this idea align with what the community values? Would they want this intervention?	Score higher if the idea feels relevant, respectful, and acceptable to the people affected.	The sticker is likely to be perceived as supportive and relevant. It is non-intrusive, easy to understand, and can be personalized to local preferences.
Ease of behaviour change	How easy will it be for the target population to adopt the intended behaviour?	Score higher if the idea reduces friction, feels simple to act on, or fits into existing routines.	It reduces the chance of forgetting by reinforcing an already intended behaviour. It fits easily into the caregiver's routine and environment.
Impact of behaviour change	If the behaviour changes, will it meaningfully influence the desired outcome?	Score higher for ideas that directly affect key health, social, or behavioural outcomes.	Timely attendance at scheduled appointments directly supports full immunization, improving health outcomes.
Measurability	Can we reliably measure whether this idea is working?	Score higher if the idea includes clear outputs, observable behaviours, or simple metrics.	Impact can be tracked by comparing clinic attendance records for caregivers who receive the sticker versus those who do not.
Cost	Is the idea relatively affordable to implement?	Score higher if it uses existing systems, requires minimal resources, or can be adapted cost-effectively.	The intervention is relatively low-cost and simple to implement using existing staff and systems.
Scalability	Could this idea be expanded to other people, settings, or regions if successful?	Score higher if it is modular, adaptable, or builds on systems that already exist.	The design is adaptable and can be scaled across different locations with minor contextual adjustments such as local language or imagery.

b. Assign a score to each idea and briefly note the rationale in the "Why?" column to document decisions and support team alignment.

c. Calculate the total score out of 30 by adding the points across each row. These totals provide a quick snapshot for internal comparison and discussion, helping the team to identify the most promising ideas. However, the scores are not final decisions or endorsements. Before sharing externally (e.g., with the Ministry of Health), ensure prioritized ideas are fully developed and contextually appropriate.

d. After the discussion, select 1–3 ideas to move forward into the next sub-step: ethics assessment.

2. Ensure the idea is ethical and inclusive

Use the Ethics Checklist worksheet to assess each prioritized idea against key ethical considerations:

a. Start with one of the prioritized intervention ideas from the previous step.

b. Review each prioritized idea and assess whether it meets the ethical standard. For each criterion, indicate one of the following:

- **Yes**: The idea clearly meets the ethical standard.
- **Yes, with potential risks to keep in mind**: The idea generally meets the standard, but carries some ethical considerations that should be monitored or addressed.
- **No**: The idea does not meet the ethical standard, and may need to be revised or reconsidered.

The table below provides guiding questions and prompts to help assess each criterion thoughtfully and consistently.

TABLE 6. ETHICS CRITERIA, PROMPT QUESTIONS AND FURTHER GUIDANCE

CRITERIA	PROMPT	GUIDANCE
Equity and Inclusion	Could this idea worsen existing inequalities or exclude certain groups?	Review whether the idea is equally accessible to people across different income levels, ethnicities, education levels, and abilities. For example, a smartphone-based intervention may exclude caregivers without digital access.
Gender and Power Dynamics	Does this idea reinforce gender stereotypes or imbalances in decision-making?	Consider whether the idea assumes who makes health decisions in the household. For example, targeting only mothers for child health may reinforce norms that exclude fathers or other caregivers.
Intersectionality	Has the idea considered how overlapping identities may shape barriers or experiences?	Reflect on how people may face compounded disadvantages (e.g. young, rural, low-income mothers). If an idea overlooks these intersecting realities, it may unintentionally ignore or marginalize key populations.
Respect for Autonomy and Consent	Does the idea allow people to make informed, voluntary decisions?	Ensure the intervention is not overly coercive or manipulative. For example, social pressure messages should not shame or pressure people into action. Information should be accurate and consent should be respected.
Protection of the Most Vulnerable	Could this idea unintentionally cause harm or increase vulnerability?	Ask whether the idea may put certain individuals at risk. For example, overburdening already stretched community workers, or exposing caregivers to stigma. Ideas should do no harm and ideally reduce vulnerability.
Transparency and Trustworthiness	Is it clear who is behind the intervention and what its purpose is?	People should understand why they are receiving a message or being asked to act. For instance, a reminder message should clearly state that it is from the Ministry of Health or another trusted source, not an anonymous sender.

Note: This ethics checklist is not drawn from a single source. It was developed for this toolkit by synthesizing key principles from UNICEF equity and safeguarding standards, WHO "Do No Harm" guidance, and global HCD practice. The aim is to provide a practical, design-focused ethics lens to help teams assess intervention ideas fairly and inclusively.

c. Discuss and document reflections. Note where the intervention feels strong, and where it may need adjustment. Capture any actions needed to improve the idea's ethical integrity in the column "Risks and adjustments needed" of the worksheet.

d. If the intervention raises ethical risks or red flags, this is the moment to rethink its design. Revisions might involve changing the language, delivery method, framing, or format to ensure the idea remains inclusive, safe, and respectful. If multiple concerns arise and cannot be addressed through adaptation, consider either:

- Reviewing the full list of ideas to choose another alternative, or
- Seeking advice from an ethics review board before proceeding.[6]

f. Repeat this process for each prioritized intervention idea.

Keep the following tip in mind when prioritizing ideas:

→ **Consider whether this idea can be tested both simply and soon.** Prioritize ideas that are small enough to prototype, but meaningful enough to learn from. If it's too complex to test quickly, consider whether it can be broken into smaller parts.

6 For more information, consult UNICEF's internal ethical review policy

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

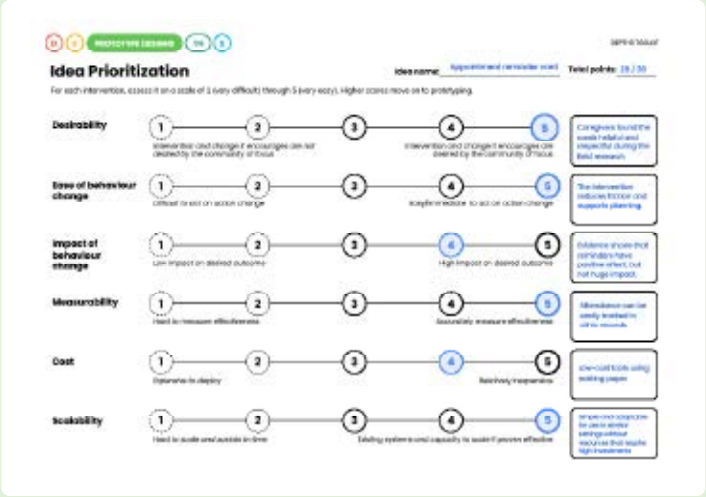
This Idea Prioritization worksheet and Ethics Checklist were not developed by the original project team. They are recreated examples based on real project data and context.

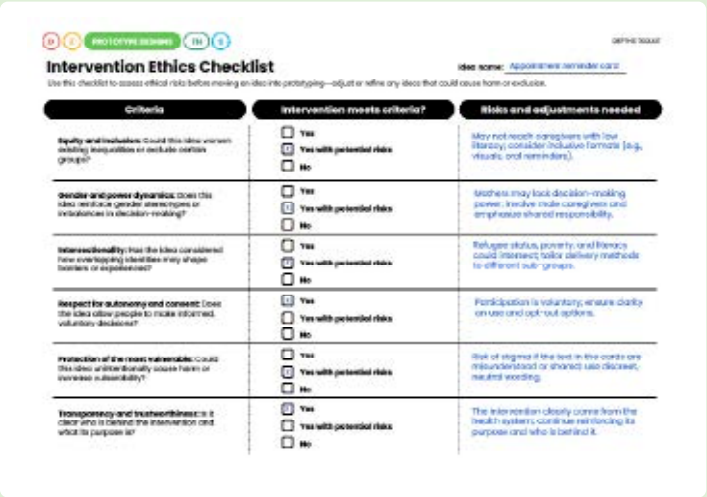
After generating and refining a diverse set of intervention ideas, the team used six criteria: desirability, ease of behaviour change, impact, measurability, cost, and scalability. These criteria helped the team to assess not only how well an idea aligned with the community's needs and preferences, but also whether it could drive meaningful change in a feasible, sustainable way.

Among the shortlisted ideas, one stood out: providing caregivers with a simple appointment card at the clinic with a reminder message of their child's next vaccination. The team rated this intervention highly on **ease** (past evidence on similar interventions showed an effective behaviour change), **cost** (it could be layered onto existing health systems), and **scalability** (it had potential to expand across regions with minimal adaptation). It was also seen as relevant and respectful since caregivers were already used to receiving information this way, and qualitative research suggested they valued simple, visual cues to remember key appointments.

Before finalizing the decision, the team applied an ethical lens, considering equity, inclusion, gender dynamics, and protection of vulnerable groups. While the intervention generally met ethical standards, the team noted a few potential risks: some caregivers might not read, leading to exclusion, and female caregivers might lack decision-making authority. As a result, small but important adjustments were made, such as designing appointment cards with visual cues to support low-literacy caregivers and considering messaging that could be discreet yet encouraging.

With the ethical review complete and behavioural promise confirmed, the team selected the appointment card to move into the next phase: prototyping and testing with caregivers.





STEP 3:

Prototype, test, and iterate interventions

In this step:

Selected ideas are brought to life by developing prototypes and by testing them with users to gather feedback and inform further refinement.

Associated tools:

- Prototype Outline
- Prototype Discussion Guide
- Prototype Synthesis

Why it matters:

Behavioural interventions do not always perform the same way in different settings. Even well-researched ideas may be ineffective or inappropriate when applied in a new context. Prototyping provides the opportunity to explore how an intervention performs, before committing significant time and resources to broader implementation. It also entails hearing the perspectives of decision-makers and those involved in implementing and delivering the intervention.

Importantly, prototyping and testing offer an early opportunity to identify challenges and unintended consequences, such as confusion, resistance, or potential harm. For example, a well-meaning reminder system might unintentionally expose private health information if messages are visible to others. Gathering feedback at this stage helps to refine core features and ensure the intervention is both effective and respectful before moving to implementation.

How to do it:

Use the Prototype Outline worksheet to further develop the selected intervention and define what form the prototype will take.

1. Describe the idea clearly

a. Start with writing a short headline and one-sentence description that explains what the intervention is and what it aims to achieve, without jargon. This helps to ensure the idea is easily understood by all stakeholders.

b. Next, define the expected reach. Specify the geographical scope (e.g., local, regional, national), including relevant locations or administrative areas. Clarify whether it targets urban, peri-urban, or rural settings, and note any unique constraints or opportunities. Estimate the number of people the intervention will reach, both directly and indirectly.

2. Formulate the intervention hypothesis

The intervention hypothesis ensures that the prototype is informed by behavioural science and focused on solving the right problem.

a. Begin by reviewing the key outputs from the **Explore & Diagnose** phase, especially the Behavioural Map and Diagnosis and the intervention ideas from Step 1 of this phase. The first task is to clearly define the target micro-behaviour that the intervention seeks to influence. Specify the intended behaviour and what success looks like.

For example: **"Increase timely completion of routine childhood vaccinations among under-vaccinated children in areas targeted by Lebanon's Accelerated Immunization Activities (AIA) programme."**

Revisit the prioritized barriers to ensure refined ideas remain anchored to the most critical ones. Reframe them briefly through COM-B to confirm alignment with behavioural domains (capability; opportunity; motivation). This restatement strengthens the intervention hypothesis by linking ideas to behavioural evidence.

b. Reflect on how the prototype is expected to shift behaviour, identifying the specific behavioural mechanisms being activated. What exactly about the prototype is expected to work, and why? Consider the following questions:

- What behavioural levers does this idea pull?
- Does this idea simplify a complex decision — making it **easy**?
- Does this idea create social visibility or support — making it **social**?
- Does this idea capture attention or motivate action — making it **attractive**?
- Does this idea meet the right moment — making it **timely**?

This step is preparation for the Theory of Change. It helps to move beyond describing the prototype to articulating why it is expected to work.

c. Take a moment to consider why this idea might work well in this setting. Does this idea build on familiar routines, trusted relationships, or existing habits in the community? This quick check-in helps to ensure that the idea is not only smart on paper but makes sense in people's real lives.

3. Materialize the idea

Once the idea is clearly defined, the next step is to make it tangible:

a. A prototype is a low-cost, preliminary version of an intervention that users and stakeholders can see, hold, act out, or experience. A prototype could take many forms depending on the nature of the intervention. Table 7 outlines a non-comprehensive set of prototype types, each useful for exploring different dimensions of the intervention. Choose a format that lets people experience the idea and generate useful feedback.

For example, the appointment card was tested as a paper mock-up: a simple, non-functional version used to check layout, wording, and user reactions before full production.

(Continues on next page)

TABLE 7. TYPES OF PROTOTYPES

TYPE OF PROTOTYPE	DESCRIPTION	EXAMPLE	USEFUL FOR
Storyboards	A series of illustrated panels or sketches that show a user moving through a scenario or using a product or service.	Illustrating a caregiver's journey to a health centre after receiving a vaccination reminder.	Visualizing user journeys, identifying gaps in service experience.
Role-playing	Acting out a scenario to understand how a user might interact with a product or service.	Simulating how a health worker might deliver a new message during a household visit.	Testing emotional reactions, tone, and flow of interpersonal experiences.
Paper Mock-ups	Hand-drawn or printed paper versions of interfaces, forms, or layouts.	Designing a prototype of a new registration form for caregivers at clinics.	Testing layout preferences, wording clarity, and ease of use.
Physical Models	Tangible models made from simple materials like cardboard or clay to simulate a physical product.	Creating a model of a new waiting area layout to test navigation and comfort.	Testing navigation, user behaviour in physical spaces, and believability.
Experience Prototypes	Simulating how a person might experience the intervention in a real or semi-real environment.	Setting up a mock clinic experience to test signage and service flow.	Testing emotional engagement, environment design, and implementation details.
Concept Videos	A short video that explains how an idea works and what problem it addresses.	Demonstrating how a community referral network would support vaccine follow-ups.	Communicating complex ideas clearly and gathering feedback across stakeholders.
Service Blueprints	A map of the service delivery process, highlighting key touchpoints, roles, and backstage systems.	Visualizing the coordination between outreach teams and PHC staff in delivering services.	Understanding system-level interactions and identifying operational gaps.
SMS or Message Scripts	Draft versions of reminder messages, invitations, or follow-ups delivered via SMS, WhatsApp, or paper.	Testing tone and clarity of vaccine reminder messages with different caregiver groups.	Refining message content and timing; testing comprehension and response.

b. Prototypes can vary in level of fidelity (i.e., how closely they resemble the final version). The following table outlines levels of fidelity for different interventions.

TABLE 8. FIDELITY LEVELS OF A PROTOTYPE

PROTOTYPE FIDELITY	DESCRIPTION	EXAMPLES
Low-fidelity	A rough, conceptual representation of the idea. Drafts are hand-drawn, mocked up with basic tools, or role-played. Useful for testing broad appeal, clarity, or basic usability before investing effort.	**Programming:** A flowchart drawn on flipchart paper with sticky notes to show how a new referral pathway might work, tested with a small group of CHWs. **Digital:** A series of phone screenshots mocked up in PowerPoint or Figma showing draft SMS or WhatsApp reminder messages (not functional yet). **Training:** A draft script of a respectful counselling session read aloud by facilitators, with peers role-playing caregivers to give feedback on tone. **In-person tool:** Hand-sketched illustrations for a flipchart or counselling aid, with placeholders ("picture of caregiver here") to test resonance before commissioning real graphics.
Medium-fidelity	A more polished version that includes real content, visuals, or partial functionality. Drafts still carry "work in progress" elements but are realistic enough for users to interact with.	**Programming:** A small trial of the referral pathway in one clinic, using photocopied draft registers to test how cases are tracked across services. **Digital:** An interactive demo of a WhatsApp chatbot with a simple decision tree, tested with 10 caregivers to see if the conversation flow makes sense. **Training:** A half-day pilot of a draft training module with frontline workers, using slide templates and draft job aids to test content relevance and pacing. **In-person tool:** A laminated draft flipchart or card-sorting activity tested with a small group of caregivers, including draft visuals and language, but not the final design.
High-fidelity	A nearly final version that closely mirrors the intervention as it would be implemented. Drafts are complete in form and function, though still piloted on a smaller scale before scale-up.	**Programming:** A fully integrated referral system piloted across several districts, with finalized registers, official protocols, and supervisory tools. **Digital:** A functioning SMS reminder system linked to live clinic scheduling, piloted with real caregivers receiving and responding to messages. **Training:** The full finalized training package delivered to health workers with printed materials, digital slides, and trained facilitators, evaluated across multiple sessions. **In-person tool:** A finalized counselling flipchart or facilitation kit with professional graphics, tested in community meetings by trained staff, with structured observation of use.

4. Plan the prototyping process

a. In the **Prototype Outline** worksheet, begin by identifying and deciding with whom the idea should be tested and fill in the corresponding section.

A strong prototype gathers input from the people who matter most to the intervention's use, delivery, and long-term success. These typically fall into three key groups:

- **End-users** (e.g., caregivers, community members): Can it work in their daily lives?
- **Deliverers** (e.g., health staff, teachers): Is it practical and acceptable for them to implement?
- **Decision-makers and long-term implementers** (e.g., programme managers, policy leads): Does it align with systems and priorities for scale?

Not all groups need to be engaged at once. Early testing should prioritize end-users, who can help assess whether the idea is clear, useful, and desirable in daily life. Based on what is learned, the idea may evolve — through small changes in design, format, or delivery- to better meet people's needs. As the concept is refined, deliverers can test feasibility in service settings, while decision-makers can weigh in on system alignment and scale-up potential. For example, when testing a vaccine appointment commitment card:

- Ask end-users, such as caregivers, if it helps them to remember and feel more confident about upcoming visits
- Ask deliverers, such as nurses, if it's easy to explain and distribute during appointments
- Share early results with implementers such as immunization managers to discuss how it might be integrated into routine systems.

Test prototypes with people of different genders, ages, socioeconomic backgrounds, and languages to reflect population diversity and capture how these factors shape experience and use.

b. Next, define the geography and reach of the proposed intervention. Use the "Where must we prototype?" section of the **Prototype Outline** worksheet to identify a suitable testing location.

Choose a prototyping location that reflects real-world conditions while allowing for manageable testing and iteration. Look for a site that is:

- **Relevant to the behavioural challenge:** A location where the target behaviour occurs often and key barriers have been observed.
- **Similar to the final implementation context:** This ensures that the insights gathered during testing will apply more broadly to implementation of the intervention.
- **Logistically feasible:** Accessible to the project team, with support from local partners and existing infrastructure.

For example, if the intervention involves testing a commitment card for caregivers to confirm and remember their child's next vaccine appointment, a suitable prototyping location might be a peri-urban clinic in the priority area. Such clinics already handle high volumes of immunization visits and face challenges with caregiver follow-through, making them appropriate for observing the target behaviour. Engaging both caregivers and health workers in this setting generates relevant insights, regular feedback, and practical opportunities for iteration.

5. Determine how success will be measured

Before testing begins, define a clear and simple learning plan to assess whether the prototype is as intended. Use the "How will we know if it's working?" section of the **Prototype Outline** worksheet to document both indicators and evidence: what success looks like, and how that success will be observed.

a. Start by identifying 2–3 observable indicators that reflect the prototype's intended effect on the target micro-behaviour. Avoid vague language such as "users liked it" or "participants were confused by the message." Instead, define specific, measurable signals that show whether the prototype is prompting the desired action.

Example: If testing a commitment card for vaccine appointments, avoid vague success criteria like "caregivers understood the reminder card." Instead, consider the following:

- *"80% of caregivers correctly stated the appointment date without being prompted"*
- *"3 out of 4 caregivers signed the card and placed it in their personal bag or health booklet before leaving"*
- *"75% of caregivers report being able to explain the card's purpose in under one minute, without confusion"*

Whenever possible, compare prototype performance to baseline practice to clarify what is changing and whether those changes are meaningful.

b. The next step is to gather feedback on whether the success indicators are being met. Use a mix of verbal, nonverbal, social, and behavioural cues to assess how the prototype is performing in real-world conditions. Where possible, prioritize direct observation over self-report. Early testing isn't about proving impact but about spotting patterns and making quick adjustments before scaling. Use Table 9 below to guide what to look for, when to collect it, and how to document it:

TABLE 9. WAYS TO GATHER FEEDBACK DURING TESTING

EVIDENCE TYPE	WHAT TO LOOK FOR	WHEN TO COLLECT	HOW TO CAPTURE IT
Verbal feedback	What participants say about the prototype — what makes sense, what's confusing, what's useful or not	Immediately after use (e.g. post-interaction or exit interviews)	Use short feedback forms, informal interviews, or audio recordings. Prompt with open-ended questions like "What did you think of this?" or "Was anything unclear?"
Nonverbal reactions	Body language, hesitation, excitement, frustration — signals that may not be spoken	During interaction with the prototype	Use observation checklists or field notes to track facial expressions, pauses, tone, or fidgeting. These cues often reveal emotional responses more clearly than words.
Social cues	How group dynamics affect use — copying others, seeking validation, deferring to authority	During group or public testing situations	Assign an observer to note if participants look to peers before acting, mimic others, or ask "What are others doing?" Use team reflections or debrief sheets to record social influence patterns.
Object interactions	How people physically engage with the prototype — do they use it, discard it, ask for help?	During and immediately after prototype use	Use tally sheets or photo logs to track use patterns: Is the object handled confidently? Is it returned, reused, or thrown away? Are instructions followed?

Using the caregiver commitment card as an example, the team can assess indicators across all four evidence types:

- **Verbal feedback** – Ask caregivers: "What does the card remind you to do?"; "Was anything on the card unclear?". If many cannot explain its purpose or misinterpret it, this signals a gap in clarity or usefulness.
- **Nonverbal reactions** – Watch for: Signs of confusion (e.g., furrowed brows, hesitation to accept or sign) and confidence (e.g., nodding, smiling, quick understanding). If caregivers seem unsure, require repeated explanations, or show visible discomfort, the design may need adjustment.
- **Social cues** – In settings with multiple caregivers, note if caregivers watch others before signing or ask peers "What did you do with yours?". This may suggest the card's use isn't intuitive or that peer behaviour influences uptake.
- **Object interactions** – Observe whether caregivers: sign the card, keep it, or discard it after receiving; ask where to store it or return it to staff; or refer back to it during the visit or later appointments. These cues help to reveal whether the card is perceived as valuable or just another handout.

Document all observations (comments, expressions, peer behaviours, object use) in the "How will we know if it's working?" section of the Prototype Outline worksheet. This clarifies what the team will track to assess whether the prototype is effective or needs refinement before rollout.

6. Conduct user testing and iterate on the prototype

The next step is to test with real users in real settings.

a. Start by defining the user testing methods. Field testing can take many forms depending on the intervention, the prototype's fidelity, learning goals, and available resources. It can range from informal conversations with a few users to structured sessions with partners. A range of helpful resources exist to support this phase, such as the IDEO Org Design Kit, which provides guidance on rapid prototyping and field testing.

Using the caregiver commitment card as an example, testing could begin with informal walkthroughs, where caregivers are shown the card immediately after booking their appointment. The facilitator might ask simple, open-ended questions such as: "Would this be helpful to you?" or "What would you do with this card after leaving the clinic?" The goal is to understand first impressions and emotional reactions.

The team might then use observation (i.e. discreetly watching whether caregivers keep the card, refer to it, or ask clinic staff about it). If there are different versions of the card (e.g., one featuring a photo of the child, another with the clinic logo), parallel testing, known as A/B testing, could be used to compare preferences and engagement.

TABLE 10. USER TESTING METHODS

USER TESTING METHOD	WHAT IT IS	WHEN TO USE IT	WHAT IT HELPS TO LEARN
Informal walkthroughs	One-on-one conversations with users, showing or explaining the prototype, and asking follow-up questions.	Early stages, when testing low-fi prototypes and getting quick gut reactions.	First impressions, clarity, and emotional response.
Role plays or simulations	Acting out how the intervention would be used in context, often with implementers or end-users.	When testing interpersonal or service-based interventions (e.g. counselling, group activities).	Feasibility, flow, user experience, and delivery challenges.
Parallel testing (A/B testing)	Presenting multiple versions of a prototype to compare reactions side-by-side.	When choosing between different formats or message framings.	Preferences, usability, and comparative appeal.
Small group discussions	Guided feedback from a few users after trying or reviewing the prototype.	When deeper insights are needed into perceptions, or social norms.	Acceptability, contextual fit, improvement ideas.
Observation-only testing	Watching users interact with the prototype without prompting or guiding them.	When the prototype can be physically used, read, or navigated by the user, or to test usability in context and observe unbiased reactions.	Natural engagement, confusion, hesitation, or errors.
Lightweight pilots	Testing a simplified version of the intervention in a real-world setting over a few days or weeks.	When a high-fi prototype needs feedback in situ.	Actual use patterns, sustained engagement, and unintended effects.
Implementer dry runs	Implementers or frontline workers rehearse delivering the intervention in a mock or low-stakes setting (e.g. practice counselling session or community talk).	When testing how easy the intervention is to deliver and whether instructions or tools are clear.	Practical delivery issues, time burden, and clarity of guidance.

b. Develop tools to guide testing sessions (e.g. discussion guides, observation plans, or dry run checklists) tailored to method, depth of engagement, and prototype fidelity.

For walkthroughs, group discussions, or pilots, a structured discussion guide helps facilitators stay focused while encouraging open feedback. The Prototype Discussion Guide worksheet provides a useful starting point, with prompts such as:

- **Introducing the purpose of the test and setting expectations:** Welcome participants, explain the session, emphasize feedback is valued and confidential, and seek consent.
- **Identifying points of confusion, resistance, or friction:** Revisit previously identified behavioural barriers and confirm whether they still feel relevant. Surface any new challenges that may have emerged.
- **Walking users through the prototype and gathering reactions (e.g., usefulness, clarity, fit):** Walk through each prototype, gather initial reactions, and explore clarity, usefulness, concerns, and suggestions for improvement.
- **Surfacing ideas for improvement or refinement:** Prompt participants to reflect on which ideas are most and least useful, and what could be changed to improve them.

This worksheet does not need to be used in its entirety or exactly as-is. Instead, adapt it by selecting the most relevant questions for the prototype method and learning goals.

For observational methods, such as watching users interact with a prototype or conducting implementer dry runs, the same core questions can be translated into simple note-taking templates or observation checklists. These help to systematically capture key behaviours and reactions, even when participants aren't speaking directly.

c. To ensure ethical and effective testing, consider the following when rolling out the user testing:

- Obtain informed consent from all participants using a simple informed consent form or verbal script (depending on context and literacy levels) to formalize this process.
- Avoid putting people in situations where they may feel judged, embarrassed, or coerced. To reduce this risk, keep these suggestions in mind:
    - Avoid testing in group settings if the prototype involves sensitive topics or personal experiences
    - Maintain a sense of neutrality in facilitation, by avoiding praising or correcting participant behaviour
    - Design neutral scripts and tools that don't imply a "right" or "wrong" response
    - Let participants interact freely, without pressure to perform a certain way
    - Avoid recording names or identifiable information, unless absolutely necessary
- Ensure that the prototype is safe and appropriate for the context.
- Debrief as a team immediately after testing to capture fresh insights.

d. After each round of user testing, synthesise what was learned. Use the Prototyping Synthesis worksheet to document the main takeaways.

Start with raw notes or observation logs from the test session(s). Review what people did, said, or showed through body language. Next, reflect on the overall direction of the intervention and respond to the following questions:

- What shows the idea aligns with its original purpose? Reflect on whether users understood the idea in the way it was intended. This helps to assess if the prototype is delivering on the behavioural hypothesis.
- What suggests the idea could lead to the intended behaviour change? Look for signals in what users say about ease, usefulness, or the likelihood of follow-through.
- What indicates the idea is addressing a specific barrier? Capture any evidence that the idea is helping to overcome the identified COM-B barrier: making something clearer, easier, more motivating, or more accessible.

At this stage, answers are tentative. Rapid testing can show whether an idea is promising or problematic, but only rigorous evaluation can confirm impact

e. Next, fill out the key changes that emerged from prototyping:

- **Keep:** Elements that worked well and should remain unchanged
- **Improve:** Elements that caused confusion, hesitation, or friction
- **Drop:** Elements that didn't land, were misunderstood, or introduced risks
- **Add:** Missing elements or ideas that could strengthen the intervention

f. Next, iterate on the prototype to refine it. Iteration is the step where the prototype is improved based on what was learned during testing. This could mean rewording messages, changing visuals, adjusting how it's introduced, or tweaking delivery materials or scripts. In some cases, it may also mean dropping an idea altogether if it doesn't perform well, or creates confusion, discomfort, or ethical concerns.

g. Once the revised version is ready, return to the field for another round of testing. There is no universal number of testing rounds that applies to all prototypes. However, at least two rounds of user testing are usually recommended, as outlined below:

- Initial testing of the early prototype to assess usability, relevance, and initial reactions.
- Follow-up testing after revisions to confirm whether changes improved performance and resolved earlier issues.

Additional rounds may be required depending on three criteria:

- **Confidence in performance**: If users still struggle to understand, accept, or use the prototype as intended, continue testing until feedback is consistently positive.
- **Complexity of the prototype**: The more complex the prototype, the more testing cycles are likely needed.
- **Range of use cases**: If the prototype will be used across multiple contexts or audience segments, test with representatives from each group.

h. Document each version, the changes made, and their impact. This informs future adaptation or scale-up. Each round can be small (5–10 users per segment), as long as sessions are focused and well-facilitated. Continue testing until repeated sessions stop generating insights.

Consider the following tips when prototyping and user testing ideas:

- **Have a "Prototype teammate" to observe what may be missed.** Even the best facilitators miss things in the moment. Assign a team member as a silent observer during testing sessions. Their only job is to take notes on non-verbal cues (e.g., hesitation, confusion, surprise) and moments when users deviate from expectations. This adds valuable perspective during synthesis, especially on usability and friction points.
- **Label each prototype version with a unique code.** Avoid confusion when iterating. Use a versioning system (e.g., "CHW job aid v2.1 – photo icon layout") on every prototype. Include the version and the key change it contains. This makes it easy to track what was tested, how it performed, and what changed in the next round.
- **Always test how it's introduced, not just the tool itself.** A prototype's success often hinges on how it's explained, handed over, or framed by implementers. During testing, observe not only how users react to the tool, but also how it's introduced. If frontline workers are unsure how to explain it, the tool may never land. Treat the handover moment as part of the prototype, and test it just as rigorously.
- **Validate the quieter voices.** Outspoken participants often dominate group testing. To ensure everyone is heard, include a "silent feedback" moment- for example, ask participants to mark parts of a paper prototype they like or dislike using stickers or symbols, before discussion starts.

CASE STUDY

Increasing childhood vaccination uptake in Lebanon

The Prototype Outline, Prototype Discussion Guide, and Prototype Synthesis were not developed by the original project team. They are recreated examples based on real project data and context.

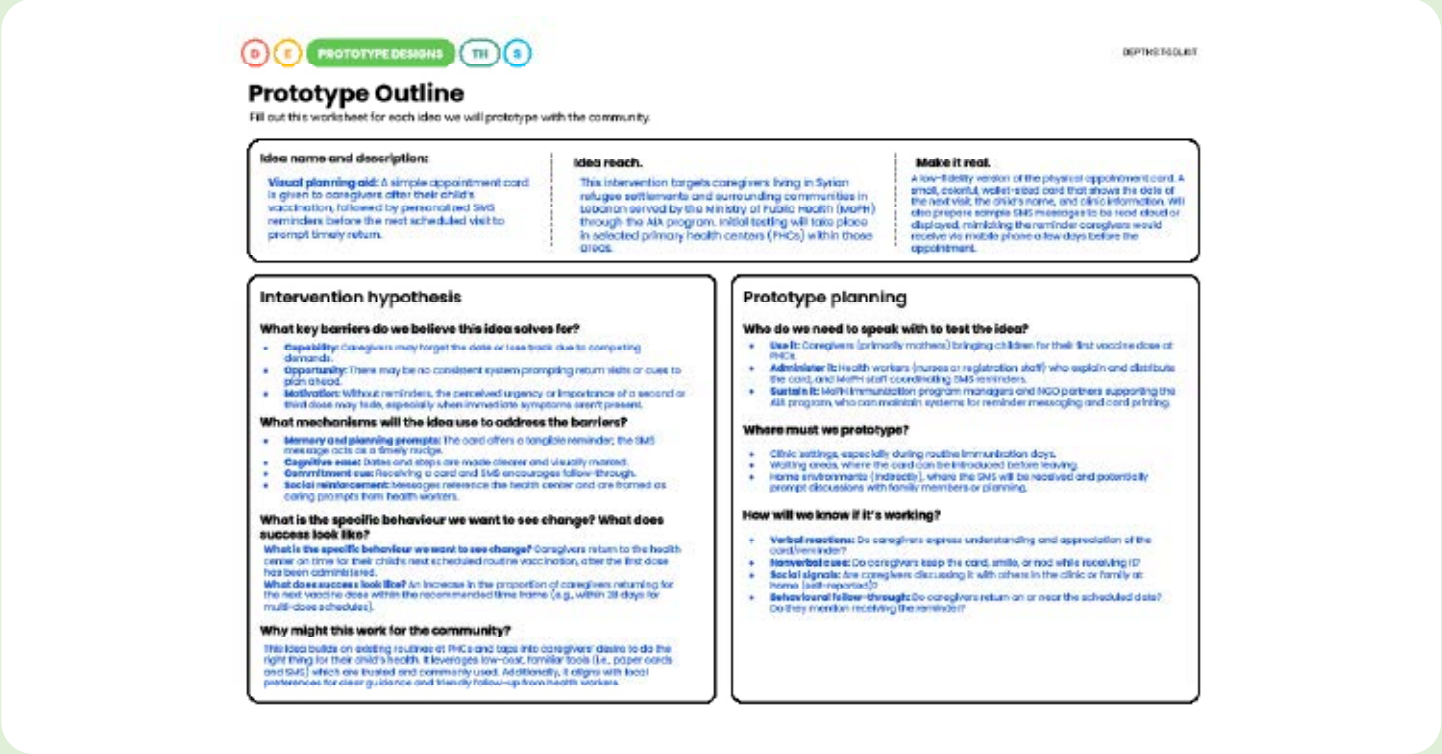
After shortlisting intervention ideas using prioritization and ethical criteria, the research team in Lebanon moved into the prototype phase. The idea selected for development was a paper-based appointment card, a simple tool designed to help caregivers remember their child's next vaccination date and return on time.

Developing the prototype

To begin, the team created a prototype outline. The intervention aimed to address specific barriers identified among caregivers in the AIA programme, particularly forgetfulness, cognitive overload, and a lack of clear guidance on when to return for the next dose.

The prototype outline clarified the intended behaviour (caregivers returning on time for the next vaccination), the expected mechanisms (enhancing memory and planning, creating a physical cue), and the motivation for its likely success (it was simple, tangible, and could fit naturally into existing caregiver routines). The prototype's reach was defined as health centres serving Syrian and Lebanese families across low-resource settings.

The team then developed a low-fidelity version of the appointment card (see on the right). This minimalist design was intentional, as it was easy to reproduce, inexpensive, and allowed the team to test core elements before investing in a more polished version.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Prototype Outline

Fill out this worksheet for each idea we will prototype with the community.

Idea name and description:

Visual planning aid: A simple appointment card is given to caregivers after their child's vaccination, followed by personalized SMS reminders before the next scheduled visit to prompt timely return.

Idea reach.

This intervention targets caregivers living in Syrian refugee settlements and surrounding communities in Lebanon served by the Ministry of Public Health (MoPH) through the AIA program. Initial testing will take place in selected primary health centers (PHCs) within those areas.

Make it real.

A low-fidelity version of the physical appointment card. A small, colorful, wallet-sized card that shows the date of the next visit, the child's name, and clinic information. Will also prepare sample SMS messages to be read aloud or displayed, mimicking the reminder caregivers would receive via mobile phone a few days before the appointment.

Intervention hypothesis

What key barriers do we believe this idea solves for?

- **Capability:** Caregivers may forget the date or lose track due to competing demands.
- **Opportunity:** There may be no consistent system prompting return visits or cues to plan ahead.
- **Motivation:** Without reminders, the perceived urgency or importance of a second or third dose may fade, especially when immediate symptoms aren't present.

What mechanisms will the idea use to address the barriers?

- **Memory and planning prompts:** The card offers a tangible reminder; the SMS message acts as a timely nudge.
- **Cognitive ease:** Dates and steps are made clearer and visually marked.
- **Commitment cue:** Receiving a card and SMS encourages follow-through.
- **Social reinforcement:** Messages reference the health center and are framed as caring prompts from health workers.

What is the specific behaviour we want to see change? What does success look like?

What is the specific behaviour we want to see change? Caregivers return to the health center on time for their child's next scheduled routine vaccination, after the first dose has been administered.

What does success look like? An increase in the proportion of caregivers returning for the next vaccine dose within the recommended time frame (e.g., within 28 days for multi-dose schedules).

Why might this work for the community?

This idea builds on existing routines at PHCs and taps into caregivers' desire to do the right thing for their child's health. It leverages low-cost, familiar tools (i.e., paper cards and SMS) which are trusted and commonly used. Additionally, it aligns with local preferences for clear guidance and friendly follow-up from health workers.

Prototype planning

Who do we need to speak with to test the idea?

- **Use it:** Caregivers (primarily mothers) bringing children for their first vaccine dose at PHCs.
- **Administer it:** Health workers (nurses or registration staff) who explain and distribute the card, and MoPH staff coordinating SMS reminders.
- **Sustain it:** MoPH Immunization program managers and NGO partners supporting the AIA program, who can maintain systems for reminder messaging and card printing.

Where must we prototype?

- Clinic settings, especially during routine immunization days.
- Waiting areas, where the card can be introduced before leaving.
- Home environments (indirectly), where the SMS will be received and potentially prompt discussions with family members or planning.

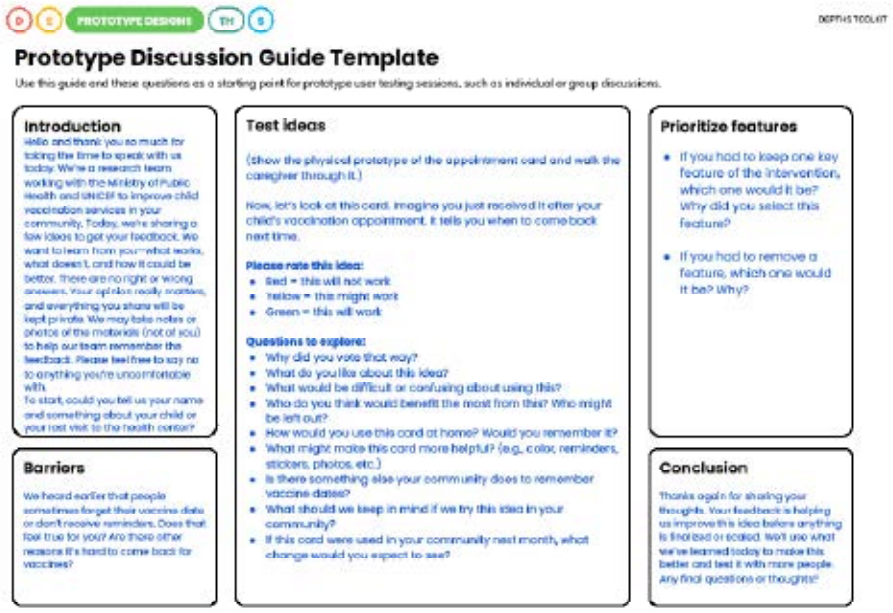
How will we know if it's working?

- **Verbal reactions:** Do caregivers express understanding and appreciation of the card/reminder?
- **Nonverbal cues:** Do caregivers keep the card, smile, or nod while receiving it?
- **Social signals:** Are caregivers discussing it with others in the clinic or family at home (self-reported)?
- **Behavioral follow-through:** Do caregivers return on or near the scheduled date? Do they mention receiving the reminder?

User testing the prototype

To ensure the card was usable and acceptable, the team conducted informal walkthroughs with caregivers in simulated clinic settings. Using a structured guide, they captured first impressions, emotional responses, and concerns.

Caregivers rated the prototype (green = works, yellow = might work, red = won't work) and explained why. Conversations explored what they liked, what confused them, who the card would help, and how it could be improved.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Prototype Discussion Guide Template

Use this guide and these questions as a starting point for prototype user testing sessions, such as individual or group discussions.

Introduction

Hello and thank you so much for taking the time to speak with us today. We're a research team working with the Ministry of Public Health and UNICEF to improve child vaccination services in your community. Today, we're sharing a few ideas to get your feedback. We want to learn from you—what works, what doesn't, and how it could be better. There are no right or wrong answers. Your opinion really matters, and everything you share will be kept private. We may take notes or photos of the materials (not of you) to help our team remember the feedback. Please feel free to say no to anything you're uncomfortable with.

To start, could you tell us your name and something about your child or your last visit to the health center?

Test ideas

(Show the physical prototype of the appointment card and walk the caregiver through it.)

Now, let's look at this card. Imagine you just received it after your child's vaccination appointment. It tells you when to come back next time.

Please rate this idea:

- Red = this will not work
- Yellow = this might work
- Green = this will work

Questions to explore:

- Why did you vote that way?
- What do you like about this idea?
- What would be difficult or confusing about using this?
- Who do you think would benefit the most from this? Who might be left out?
- How would you use this card at home? Would you remember it?
- What might make this card more helpful? (e.g., color, reminders, stickers, photos, etc.)
- Is there something else your community does to remember vaccine dates?
- What should we keep in mind if we try this idea in your community?
- If this card were used in your community next month, what change would you expect to see?

Prioritize features

- If you had to keep one key feature of the intervention, which one would it be? Why did you select this feature?
- If you had to remove a feature, which one would it be? Why?

Barriers

We heard earlier that people sometimes forget their vaccine date or don't receive reminders. Does that feel true for you? Are there other reasons it's hard to come back for vaccines?

Conclusion

Thanks again for sharing your thoughts. Your feedback is helping us improve this idea before anything is finalized or scaled. We'll use what we've learned today to make this better and test it with more people. Any final questions or thoughts?

Key feedback and iterations

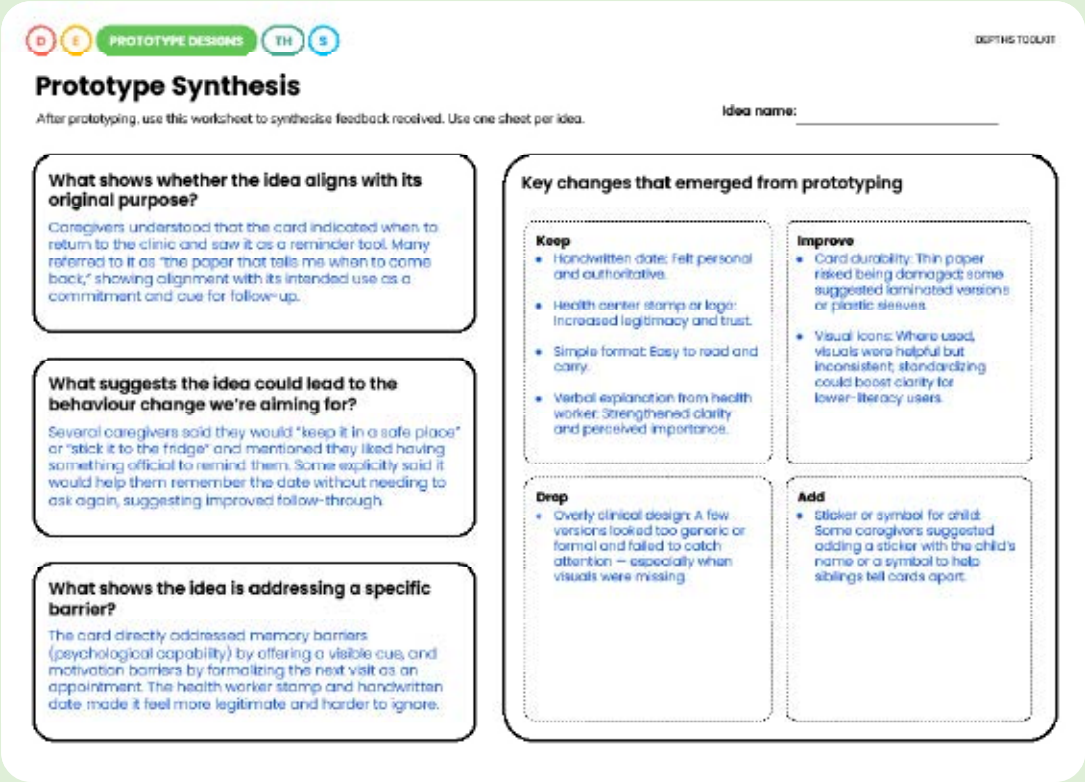
Feedback was largely positive. Caregivers valued having the date written down and said it created a sense of "appointment obligation," making the visit feel more official.

At the same time, the testing revealed specific opportunities for refinement:

- **Keep:** handwritten date, health centre stamp/logo, simple layout, verbal explanation from staff
- **Improve:** card durability (lamination/sleeves), standardize icons for low-literacy users
- **Drop:** overly clinical or generic designs that felt impersonal and easy to ignore
- **Add:** stickers or child symbols to differentiate siblings' cards and engage children

This feedback was documented using the **Prototyping Synthesis** worksheet and directly informed improvements to the card.

These changes helped increase the card's relevance, usability, and resonance with both caregivers and health workers, building a stronger foundation for testing at scale.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Prototype Synthesis

After prototyping, use this worksheet to synthesise feedback received. Use one sheet per idea.

Idea name: _____

What shows whether the idea aligns with its original purpose?

Caregivers understood that the card indicated when to return to the clinic and saw it as a reminder tool. Many referred to it as "the paper that tells me when to come back," showing alignment with its intended use as a commitment and cue for follow-up.

What suggests the idea could lead to the behaviour change we're aiming for?

Several caregivers said they would "keep it in a safe place" or "stick it to the fridge" and mentioned they liked having something official to remind them. Some explicitly said it would help them remember the date without needing to ask again, suggesting improved follow-through.

What shows the idea is addressing a specific barrier?

The card directly addressed memory barriers (psychological capability) by offering a visible cue, and motivation barriers by formalizing the next visit as an appointment. The health worker stamp and handwritten date made it feel more legitimate and harder to ignore.

Key changes that emerged from prototyping

Keep
- Handwritten date: Felt personal and authoritative.
- Health center stamp or logo: Increased legitimacy and trust.
- Simple format: Easy to read and carry.
- Verbal explanation from health worker: Strengthened clarity and perceived importance.

Improve
- Card durability: Thin paper risked being damaged; some suggested laminated versions or plastic sleeves.
- Visual icons: Where used, visuals were helpful but inconsistent; standardizing could boost clarity for lower-literacy users.

Drop
- Overly clinical design: A few versions looked too generic or formal and failed to catch attention — especially when visuals were missing.

Add
- Sticker or symbol for child: Some caregivers suggested adding a sticker with the child's name or a symbol to help siblings tell cards apart.

STEP 4:
Theory of Change

In this step:

As the process moves from prototype to evaluating a pilot, it is essential to clarify the logic behind the intervention. How is the intervention expected to shift behaviour, and contribute to the broader outcome articulated during the Define phase? This requires a practical and evidence-informed Theory of Change (ToC), which includes the following:

- actors involved
- needs and barriers identified
- proposed activities
- how these factors are expected to produce changes in behaviour and outcomes

The Theory of Change serves two core functions:

- **Internal clarity:** Ensuring alignment within the team on how the intervention works and what success looks like
- **External alignment:** Helping funders, partners, and decision-makers quickly grasp the rationale and pathway to impact

Associated tools:

- Theory of Change

Why it matters:

A ToC is a practical tool that helps teams to clearly map out **how** an intervention is expected to create change — inking activities to targeted behaviours and intended outcomes. Done well, the ToC tells us the mechanisms behind change. A strong ToC communicates:

- The specific behaviour being shifted
- Why this intervention will work in this context —the mechanisms it activates, such as trust, self-efficacy, or social influence
- What assumptions are being made and where the risks are

A ToC helps to move beyond a checklist of activities — serving a purpose beyond donor reporting. Instead of simply linking inputs to outputs, it lays out the logic of change: what needs to shift, in what order, under what conditions, and through which underlying psychological or behavioural mechanisms. This helps avoid "mechanical" design that skips over real drivers of behaviour.

Example of a weak or "mechanical" ToC: If we train community health workers, then immunization rates will increase. This skips over the "how." It does not specify which behaviours need to change (e.g., caregivers overcoming hesitancy, CHWs consistently conducting follow-ups), nor does it identify the mechanisms that training is meant to activate (e.g., building trust, shifting norms, improving self-efficacy). Without this detail, the causal pathway is incomplete and risks overlooking the real bottlenecks to behaviour change.

Example of a stronger behavioural ToC: If we equip CHWs with empathetic listening techniques and tools to address caregivers' fears about side effects (activity), then caregivers will feel more understood and confident in their decision (mechanism: building trust and self-efficacy). With regular, trusted follow-up through community visits (activity), caregivers will also perceive vaccination as a community norm (mechanism: social influence). Together, these shifts increase the likelihood that caregivers return for all doses (behaviour), leading to higher completion of immunization schedules (outcome).

Many interventions succeed only if prerequisites are in place (e.g., trust built first, reliable resources, social approval). A good ToC makes these dependencies explicit so the logic reflects how change is likely to happen in practice.

The ToC is a living document that can be revisited and adapted over time. It provides the foundation for monitoring and evaluation in later phases, and helps spot gaps, risks, or ethical concerns before they emerge. It creates a shared map of how change is expected to happen, making assumptions explicit and clarifying pathways for the team and partners.

How to do it:

Each of the following sub-sections corresponds to a component of the Theory of Change worksheet, which helps to formalize the core logic of the intervention and prepares it for piloting or evaluation.

1. Identify the actor

a. Start with the first column on the Theory of Change worksheet by identifying the key actors — the individuals, groups, or institutions whose involvement will determine whether the intervention succeeds. This includes:

- **End-users:** Those who will directly experience the intervention (e.g., caregivers, adolescents, community members)
- **Deliverers:** Those responsible for carrying out the intervention on the ground (e.g., health workers, outreach staff, educators)
- **Implementers or decision-makers:** Those who influence, approve, or sustain the intervention over time (e.g., programme managers, local leaders, government officials)

b. For each actor, go beyond listing names or roles, asking "What does this actor need to do differently for the intervention to succeed?" For example:

- End-users, such as caregivers, may need to act on follow-up reminders and attend scheduled visits
- Deliverers, such as health workers, may need to introduce new tools or adopt a different counselling approach
- Implementers, such as programme managers, may need to allocate time, staff, or political will to support the intervention at scale

c. Note any actors who may block or slow progress. Consider:

- Are there individuals with veto power who are not yet supportive?
- Do any actors have conflicting priorities or incentives?

2. Define needs and key barriers

Next, use the second column of the worksheet to capture the behavioural barriers that the intervention is designed to address. This section should draw from the behavioural mapping and diagnosis completed during the **Explore & Diagnose** phase.

The Prototype Outline also contains a summary of those barriers — updated prior to testing — and can serve as a primary starting point. Revisit the needs and barriers documented, refining the barriers based on what was learned during prototype and testing. Have any become more or less important based on user feedback? Did any new barriers emerge during testing?

Record these key barriers in the second column of the worksheet. Focus only on the barriers that this intervention is directly trying to address, rather than the full set identified during earlier research. To structure the barriers in this section, apply the COM-B framework:

- **Capability:** e.g. gaps in knowledge, limited confidence, or unclear information
- **Opportunity:** e.g. time constraints, geographic access, social influence, or service availability
- **Motivation:** e.g. underlying beliefs, emotions, priorities, or habits

3. Clarify resources and implementation conditions

List the resources and conditions needed for the intervention to succeed in the third column of the worksheet. These may include:

- **Human resources:** e.g., trained health workers, peer mobilizers, supervisors
- **Materials and tools:** e.g., commitment cards, mobile phones, printed job
- **Delivery requirements:** e.g., airtime or data subsidies, transport for outreach staff
- **Enabling factors:** e.g., political support, integration with existing programs, community trust in frontline workers, or heightened public urgency due to a recent outbreak

Be sure to flag any critical dependencies or risks. For example:

- "This approach only works if mobile data is subsidized."
- "In order to use CHW time, Ministry approval is required."
- "Community gatekeepers must approve the intervention."

While some of these elements may have been noted and discussed earlier, this is the moment to consolidate them in one place. Capturing them clearly ensures that everyone understands the conditions needed for success, along with highlighting what may need to be secured before moving toward scale or evaluation.

4. Outline the intervention activities

In the fourth column of the worksheet, describe the specific actions your team will implement to address the priority barriers identified. They form the operational core of your intervention and set the stage for both implementation and evaluation. Each activity should be:

- Directly linked to a behavioural barrier
- Explicitly tied to a mechanism of change (e.g., reminders, planning prompts, trusted messengers, simplification, incentives). For example:
    - If caregivers forget appointments (capability barrier), the activity might be sending SMS reminders tied to clinic dates (mechanism: salience + memory aid).
    - If CHWs feel unsure how to address hesitancy (motivation + capability barriers), the activity might be interactive training with role-play and tailored job aids (mechanism: self-efficacy + trusted messenger).
    - If clinic flow is overwhelming (opportunity barrier), the activity could be restructuring the waiting area to allow for quieter one-on-one discussions (mechanism: reducing friction + enabling supportive environment).

This section should explain not only what will be done, but why it's expected to influence behaviour, grounding each activity in both practical logic and behavioural insight. Use this checklist to pressure-test your activities:

- Is each activity clearly tied to a specific barrier?
- Do we understand the behavioural mechanism it activates?
- Is it feasible with current resources and systems?
- Is there any evidence (local or global) that it could work?
- Is it acceptable to those delivering and receiving it?

5. Psychological/Behavioural Mechanisms

The fifth column is the core of the Theory of Change in a behavioural science project. It explains **why** the intervention is expected to work, by identifying the underlying psychological or behavioural mechanisms that act as levers of change. These mechanisms connect the designed activities to the expected shifts in behaviour, making the causal pathway both clear and testable. Without this step, the Theory of Change risks becoming a standard input–output chain, rather than a behavioural science tool.

Each activity should be tied to one or more mechanisms of change, such as:

- **Self-efficacy**: increasing people's confidence that they can perform the behaviour
- **Social norms**: signalling that others in their community value or practice the behaviour

- **Habit formation**: reinforcing small, repeatable actions until they become automatic
- **Trust**: strengthening belief in the credibility and intentions of the messenger or system
- **Reduced cognitive load**: making the desired behaviour easier by simplifying choices or reducing mental effort
- **Salience**: drawing attention to the behaviour at the right moment, making it difficult to ignore

For example:

- SMS reminders work because they reduce cognitive load (people don't have to remember the date themselves) and increase salience (the message arrives close to the appointment).
- Interactive CHW training strengthens trust (caregivers perceive health workers as credible) and builds self-efficacy (CHWs feel confident responding to hesitancy).
- Restructuring clinic flow reduces friction (opportunity barrier) and creates supportive one-on-one interactions, which foster trust and social support.

These mechanisms also provide a clear basis for testing during piloting and later evaluation, making it possible to assess whether they were activated as intended; if not, the design can be refined.

6. Define the outputs

In the sixth column of the worksheet, list the immediate, observable results of the intervention activities. These are not behaviour changes but the short-term signals that show whether the intervention is being delivered and used as intended, providing early clues that it may be on the right track. Outputs show that the foundational elements of the intervention (e.g., tools, messages, and training) are reaching the right people in the right way.

Where the previous column (Psychological/Behavioural Mechanisms) explained why an activity is expected to work, this column captures what can be directly observed or counted to confirm the activity is happening as designed. Good outputs meet the following criteria:

- **Observable**: they can be seen or tracked directly
- **Quantifiable**: they can be counted or recorded reliably with basic metrics
- **Logically linked**: they are the direct result of an intervention activity

Examples include caregivers receiving and holding reminder cards, health workers using new scripts during counselling sessions, or volunteers distributing planning tools during home visits. These outputs build on the early indicators of success documented in the Prototype Outline, but are now positioned as part of the overall change pathway.

Outputs are generally measured through tools such as observation checklists, distribution logs, training attendance sheets, or implementer self-reports to document who received what, when, and how consistently. For example, observation can confirm whether a caregiver keeps a reminder card; checklists can track how many home visits included the new tool.

It's important to avoid inventing mechanisms or assuming that change has occurred. Outputs should reflect what was actually observed during user testing or documented in past research — not what is hoped to occur. Keep the focus on immediate, concrete signals that the intervention is reaching people as intended.

Note the difference between outputs and behavioural outcomes, as detailed below :

- **Outputs** are the direct and immediate results of the intervention activities. They indicate what was delivered, to whom, and how. Outputs are observable and measurable right away — they don't require interpretation, or assume a change in behaviour, belief, or attitude.
- **Intermediate outcomes** reflect early behavioural shifts or cognitive changes that occur after exposure to the intervention. These often indicate that the intervention is beginning to influence motivation, ability, or opportunity, but they are not guaranteed just because an activity has occurred.

To clarify the distinction between the two, Table 11 below compares common intervention activities with examples of correct outputs, and incorrect statements often mistaken for outputs.

Remember: outputs answer the question, "Did this happen as a direct result of the activity, and can we see or count it right away?" In contrast, intermediate outcomes reflect early behavioural or cognitive changes that may occur after exposure to the intervention.

TABLE 11. STRONG VS. WEAK OUTPUT STATEMENTS

INTERVENTION	WEAK OUTPUTS	WHY IT'S WEAK	STRONG OUTPUT
Visual reminder card	Caregivers understood the importance of follow-up.	This is an intermediate outcome, not an immediate product of the intervention.	Caregivers leave the clinic with a tangible, visible reminder of their child's next vaccine date.
SMS reminder	SMS messages were well received by caregivers.	This is a subjective reaction that doesn't confirm the message was sent or interacted with.	Caregivers receive SMS reminders 24 hours before a scheduled visit.
Community dialogue session	Participants felt more confident discussing vaccines.	This describes a psychological shift — an intermediate outcome — not an observable result.	Community members attend a dialogue session facilitated by trained peer educators.
Job aid for health workers	Health workers provided better counselling.	This assumes an improved performance without direct observation or measurement.	Health workers use the new job aid during counselling sessions.

7. Identify intermediate outcomes

Intermediate outcomes are the short-term shifts, often cognitive or emotional, that occur after the intervention is delivered, but before the target behaviour changes. They signal whether the intervention is working and influencing the necessary levers for change, such as attention, intention, motivation, planning, or social norms. To make these outcomes practical and measurable, use observable proxies rather than generic terms. Instead of saying "motivation improved," describe what that looks like. For example:

- Caregivers bring their commitment card to the next visit (proxy for planning and follow-through).
- Participants initiate conversations about the topic with peers (proxy for increased comfort or perceived importance).
- Adolescents correctly recall the vaccination schedule after an interactive group session (proxy for improved attention and memory).

Table 12 compares common intervention activities with examples of correct intermediate outcomes, and incorrect statements often mistaken for them. Use this table to ensure that intermediate outcomes are realistic, measurable, and grounded in what can actually be observed or assessed.

TABLE 12 USEFUL VS. LESS USEFUL INTERMEDIATE OUTCOME STATEMENTS

INTERVENTION	WEAK INTERMEDIATE OUTCOMES	WHY IT'S WEAK	USEFUL INTERMEDIATE OUTCOMES
Commitment card	Caregivers receive the card.	This is an output, not a change in perception or behaviour.	Caregivers bring the card back to the next appointment (proxy for planning).
Peer group session	Participants feel more empowered.	Vague,subjective, and difficult to measure.	Participants initiate discussion about the topic with peers (proxy for confidence).
Health worker training	Providers are trained in new scripts.	This is an output, not a change in perception or behaviour.	Providers begin using new phrases when counselling (proxy for uptake of new behavior).
Reminder SMS	Caregivers are motivated to attend.	"Motivated" is abstract and unmeasurable.	Caregivers report planning transport or arranging childcare (proxy for intention and planning).

8. Write down the primary behavioural outcome

In the seventh column of the Theory of Change worksheet, record the specific behaviour the intervention ultimately aims to change. This is the specific, measurable behaviour identified during the **Define** and **Explore & Diagnose** phases. For example, this could entail:

- Attending a scheduled vaccine appointment
- Completing the full vaccine schedule
- Bringing a child to a health facility on time

The primary behavioural outcome should logically follow from the intermediate outcomes, and reflect the most critical action the intervention is designed to influence.

9. Map the intended impact

The final column outlines the broader impact the intervention seeks to achieve. This should align with the high-level Outcome Statement from the Define stage.

The impact should build directly on the previous sub- step: the primary behavioural outcome. For example, if more caregivers bring their children for scheduled vaccinations, immunization coverage improves — especially in high-risk districts. This leads to fewer missed or delayed vaccines and reduces the likelihood of outbreaks like measles or mumps, protecting both individual children and the wider population.

In this case, the behavioural outcome is increased caregiver follow-through on scheduled childhood vaccinations. This supports the broader goal of increasing the timely completion of routine childhood vaccinations among un- or under-vaccinated children in areas targeted by Lebanon's Accelerated Immunization Activities (AIA) programme. By clearly tracing the pathway from intervention to action to long-term impact, this step confirms the logic, relevance, and potential impact of the behavioural interventions.

Consider the following tips when developing a Theory of Change:

- **Use "so that..." chains to articulate links and tighten logic.** For example: We are giving CHWs counselling job aids so that they feel more confident responding to vaccine hesitancy, so that they address caregiver concerns more consistently, so that more caregivers complete the schedule. This exposes missing steps or leaps in logic.
- **Assign a "risk rating" to each ToC assumption.** After drafting the ToC, create a quick matrix where each causal link or dependency is assigned a confidence level:
    - Green→ High = Based on strong local evidence or consistent user testing
    - Yellow→ Medium = Based on global precedent, but untested locally
    - Red→ Low = Based on team hunch or optimistic assumption

    This highlights risks and guides which elements to test first in pilots or monitor most closely during rollout.

CASE STUDY

Increasing childhood vaccination uptake in Lebanon

This Theory of Change was not developed by the original project team. It's a recreated example based on real project data and context.

The team built a clear Theory of Change to ensure internal alignment and clarity on what and how the intervention was trying to achieve.

The actors included caregivers (mostly Syrian and Lebanese mothers), frontline health workers delivering the cards, and the Ministry of Public Health and AIA programme teams supporting the rollout.

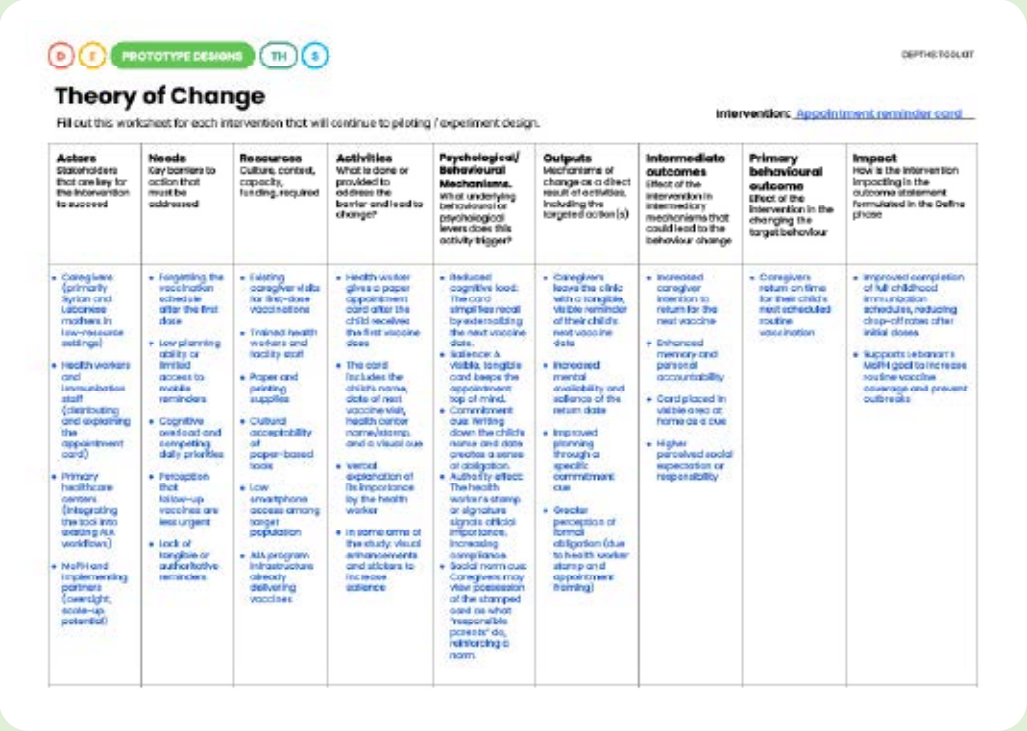
The barriers identified in previous steps were mapped in the ToC: many caregivers forgot return dates, felt overwhelmed by daily demands, or lacked reminders. Resource constraints (i.e., low smartphone penetration, limited digital infrastructure, overstretched clinics) meant a low-tech, low-cost solution was needed.

The activity itself was simple: after the first vaccine dose, health workers issued an appointment card with the next visit date. The card included visual symbols and health centre stamps to reinforce importance. The aim was not just to inform, but to nudge caregivers by embedding reminders into daily life.

- **Reduced cognitive load**: the card externalized memory, so caregivers didn't need to rely on recall alone.
- **Salience**: a visible, tangible reminder kept the date top of mind.
- **Commitment cue**: the child's name and official health worker stamp signalled obligation and formality.
- **Social norm signal**: possessing a stamped card reinforced the sense that "responsible parents" follow the schedule.

The outputs of this activity were small but powerful: caregivers left the clinic with a visible commitment device that simplified remembering the next visit. Over time, this intervention had clear intermediate outcomes: higher intention to return, increased mental recall of the date, and in some cases, even a visible display of the card in the home, serving as a continuous cue.

Ultimately, this led to the primary behavioural outcome the study aimed for: caregivers returned on time for their child's next vaccine visit. At scale, the intervention contributed to the broader impact of improving childhood immunization coverage and preventing vaccine-preventable diseases in high-risk communities.



D E PROTOTYPE DESIGNS TH S

DEPTHS TOOLKIT

Theory of Change

Fill out this worksheet for each intervention that will continue to piloting / experiment design.

Intervention: Appointment reminder card

Actors Stakeholders that are key for the intervention to succeed	**Needs** Key barriers to action that must be addressed	**Resources** Culture, context, capacity, funding, required	**Activities** What is done or provided to address the barrier and lead to change?	**Psychological/ Behavioural Mechanisms.** What underlying behavioural or psychological levers does this activity trigger?	**Outputs** Mechanisms of change as a direct result of activities, including the targeted action(s)	**Intermediate outcomes** Effect of the intervention in intermediary mechanisms that could lead to the behaviour change	**Primary behavioural outcome** Effect of the intervention in the changing the target behaviour	**Impact** How is the intervention impacting in the outcome statement formulated in the Define phase
• Caregivers (primarily Syrian and Lebanese mothers in low-resource settings) • Health workers and immunization staff (distributing and explaining the appointment card) • Primary healthcare centers (integrating the tool into existing AIA workflows) • MoPH and implementing partners (oversight, scale-up potential)	• Forgetting the vaccination schedule after the first dose • Low planning ability or limited access to mobile reminders • Cognitive overload and competing daily priorities • Perception that follow-up vaccines are less urgent • Lack of tangible or authoritative reminders	• Existing caregiver visits for first-dose vaccinations • Trained health workers and facility staff • Paper and printing supplies • Cultural acceptability of paper-based tools • Low smartphone access among target population • AIA program infrastructure already delivering vaccines	• Health worker gives a paper appointment card after the child receives the first vaccine dose • The card includes the child's name, date of next vaccine visit, health center name/stamp, and a visual cue • Verbal explanation of its importance by the health worker • In some arms of the study: visual enhancements and stickers to increase salience	• Reduced cognitive load: The card simplifies recall by externalizing the next vaccine date. • Salience: A visible, tangible card keeps the appointment top of mind. • Commitment cue: Writing down the child's name and date creates a sense of obligation. • Authority effect: The health worker's stamp or signature signals official importance, increasing compliance. • Social norm cue: Caregivers may view possession of the stamped card as what "responsible parents" do, reinforcing a norm.	• Caregivers leave the clinic with a tangible, visible reminder of their child's next vaccine date • Increased mental availability and salience of the return date • Improved planning through a specific commitment cue • Greater perception of formal obligation (due to health worker stamp and appointment framing)	• Increased caregiver intention to return for the next vaccine • Enhanced memory and personal accountability • Card placed in visible area at home as a cue • Higher perceived social expectation or responsibility	• Caregivers return on time for their child's next scheduled routine vaccination	• Improved completion of full childhood immunization schedules, reducing drop-off rates after initial doses • Supports Lebanon's MoPH goal to increase routine vaccine coverage and prevent outbreaks

Final checklist for *Prototype Designs*

- [ ] "How might we" questions
- [ ] Intervention ideas, applying the EAST framework
- [ ] Prioritized ideas using criteria and ethics considerations
- [ ] Prototype Outline worksheet
- [ ] Prototyping Synthesis worksheet after completing user testing
- [ ] Theory of Change

Optional:

- [ ] Design provocations
- [ ] Localized evidence findings *on previous interventions*

Learn more

This field guide is designed to teams with practical tools, frameworks, and methodologies to apply behavioural science to a range of real-world challenges. As behavioural science draws from multiple disciplines — including human-centred design, experimental economics, and systems thinking — we've curated a selection of approaches that reflect this diversity. The following section offers additional resources to explore specific topics introduced in the guide, along with the option to continue a self-paced learning journey.

"I want to explore different design methods for ideation and prototyping."

101 Design Methods by Illinois Institute of Technology's Institute of Design's professor, Vijay Kumar, offers a structured toolkit for innovation, helping teams move from research insights to actionable ideas through clearly defined steps. It's especially helpful if group ideation is facilitated and needs a clear structure. For hands-on prototyping advice, IDEO's Design Kit offers tools like "Determine what to prototype", which helps clarify what to test, at what level of fidelity, and with whom.

"I want to explore more creativity and brainstorming activities."

The IDEOU Brainstorming Toolkit offers practical exercises to spark ideas, encourage participation, and stretch group thinking. For a more structured approach to creative disruption, Edward de Bono's Lateral Thinking methods provide tools to challenge assumptions and generate unconventional solutions. To sharpen idea generation, the Interaction Design Foundation's guide to "How Might We" questions explains how to frame prompts that are focused, open-ended, and generative.

"I want to learn more about participatory design and community voice."

Equity-Centred Community Design by the Creative Reaction Lab offers a practical, justice-oriented lens for working with marginalized groups. Their Field Guide includes methods to reflect on power dynamics and co-create responsibly. The HCD for Health website is another great resource for those working in global health, offering tools and case studies that show how human-centred design can be embedded in field projects.

"I want to strengthen the behavioural foundations behind the interventions."

The Behavioural Drivers Model (BDM) and MINDSPACE offer practical ways to apply behavioural frameworks beyond EAST. While the first framework (BDM) is helpful to ideate interventions specifically aimed at improving immunization, the latter is a general framework that could be helpful to generate any type of idea.

"I want to better understand the role of ethics in prototyping and testing with users."

The Belmont Report remains a foundational resource for ethical principles in research with human subjects, focusing on respect, beneficence, and justice. For teams working within institutions, the CITI Programme offers accessible training modules on research ethics, informed consent, and working with vulnerable populations. For a shorter, applied overview of research ethics principles in practice, St. George's University offers a digestible summary of key topics. Finally, the Behavioural Science for Good Hub is a helpful framework to reflect whether the behavioural interventions that will be implemented are ethical.

"I want to deepen my understanding of Theory of Change and impact pathways."

The SBC Guidance website provides clear guidance on how to articulate a Theory of Change that links behaviour, evidence, and outcomes. It includes helpful prompts to connect your intervention to larger health or development goals.

Resources

1. And Academy. Guide to Ideation in Design Thinking (With Techniques). Accessed July 2025. <https://www.andacademy.com/blog/guide-to-ideation-in-design-thinking>.
2. Arouanda. The Essential Ideation Techniques for Design Thinking. Accessed July 2025. <https://arouanda.medium.com/the-essential-ideation-techniques-for-design-thinking>.
3. Behavioural Insights Team. EAST: Four Simple Ways to Apply Behavioural Insights. London: BIT, 2014. <https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights/>.
4. Behavioural Insights Team. MINDSPACE: Influencing Behaviour through Public Policy. London: BIT, 2010. <https://www.bi.team/publications/mindspace/>.
5. Bioethics Center. "CITI Training." Accessed July 2025. <https://bioethicscenter.net/citi-traininng/>.
6. Cambridge University Press. "Nudge FORGOOD." Behavioural Public Policy. Accessed July 2025. <https://www.cambridge.org/core/journals/behavioural-public-policy/article/abs/nudge-forgood/06BC9E9032521954E8325798390A998A>.
7. CITI Program. "About Us." Accessed July 2025.<https://about.citiprogram.org/>.
8. CRx Lab. Field Guide to Equity-Centered Community Design. Accessed July 2025. <https://crxlab.org/shop/p/field-guide-equity-centered-community-design>.
9. de Bono Group. "Lateral Thinking – Core Programs." Accessed July 2025. <https://www.debonogroup.com/services/core-programs/lateral-thinking/>.
10. Demand Hub. behavioural Integration for Impact (BII) Evaluation Guide. Accessed July 2025.< https://demandhub.org/bii-evaluation-guide/>.
11. Design Kit. "Determine What to Prototype." IDEO.org. Accessed July 2025. <https://www.designkit.org/methods/determine-what-to-prototype.html>.
12. Design Kit. The Human-Centered Design Toolkit. IDEO.org. Accessed July 2025. <https://www.designkit.org/resources/1.html>.
13. HCDExchange. HCD4Health. Accessed July 2025. <https://www.hcd4health.org/>.
14. Human Behaviour-Change Project. Theory and Technique Tool. Accessed July 2025.< https://theoryandtechniquetool.humanbehaviourchange.org/>.
15. IDEO. Design Kit: Methods. Accessed July 2025.< https://www.designkit.org/methods>.
16. IDEO U. "10 Idea Generation Activities for Better Ideas." Accessed July 2025. <https://www.ideou.com/blogs/inspiration/10-idea-generation-activities-for-better-ideas>.
17. IDEO U. "Brainstorming Resources." Accessed July 2025. <https://www.ideou.com/pages/brainstorming-resources>.
18. Innovation for Poverty Action. Guiding Your Program to Build a Theory of Change. Accessed July 2025. <https://www.poverty-action.org/sites/default/files/Guiding-Your-Program-to-Build-a-Theory-of-Change.pdf>.
19. Interaction Design Foundation. "How Might We." Accessed July 2025. <https://www.interaction-design.org/literature/topics/how-might-we>.
20. Interaction Design Foundation. "Prioritisation Frameworks." Accessed July 2025. <https://www.interaction-design.org/literature/topics/prioritization>.
21. International Design Foundation. "What Is Ideation?" Accessed July 2025. <https://www.interaction-design.org/literature/topics/ideation>.
22. MakeIt.Tools. "MakeIt." Accessed July 2025.< https://www.makeit.tools/>.
23. SBCC Implementation Kit. "Selecting Results." Accessed July 2025. <https://www.sbcguidance.org/create/selecting-results>.

24. St. George's University of London. "Principles of Research Ethics." Accessed July 2025. <https://www.citystgeorges.ac.uk/research/support/integrity-and-ethics/ethics/principles>.

25. Sullivan, Claire M., and Donna Spiegelman. "Implementation Science: A Cross-Disciplinary Perspective." Implementation Science 6, no. 42 (2011). <https://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-6-42>.

26. Tudor, Grant, and Benjamin Hickler. Demand for Health Services Field Guide: A Human-Centred Approach. Accessed July 2025.

27. UNICEF. Co-Creation Playbook for Community Solutions to Child Marriage and FGM. Accessed July 2025. <https://www.unicef.org/documents/co-creation-playbook-for-community-solutions>.

28. UNICEF. HCD for Health Guide. Accessed July 2025.

29. UNICEF. The Behavioural Drivers Model. Amman: UNICEF MENA, 2020. <https://www.unicef.org/mena/media/5586/file/The_Behavioural_Drivers_Model_0.pdf>.

30. UNICEF. UNICEF Publications – Latin America and the Caribbean. Accessed July 2025.< https://www.unicef.org/lac/en/publications>.

31. UNICEF ROSA. Social and Behaviour Change (SBC) Programme Guidance. Kathmandu: UNICEF ROSA, 2021.< https://www.unicef.org/rosa/media/10356/file>.

32. UNICEF. Results Framework and Theory of Change. Accessed July 2025. <https://www.unicef.org/rosa/documents/results-framework-and-theory-change>.

33. U.S. Department of Health & Human Services. The Belmont Report. Washington, DC: Office for Human Research Protections, 1979. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>.

34. Vijay Kumar. 101 Design Methods: A Structured Approach for Driving Innovation in Your Organization. Hoboken, NJ: Wiley, 2013.<http://www.101designmethods.com/>.

Test Hypotheses

Welcome to *Test Hypotheses*!

The *Test Hypotheses* phase examines whether the selected intervention contributes to measurable changes in the target behaviour by addressing key behavioural barriers (identified during the *Define* and *Explore & Diagnose* phases).

The insights generated here guide whether an intervention should be adapted, scaled, or discontinued, and lay the foundation for the Scale phase, where evidence is translated into broader action.

Unlike previous chapters that anchor on specific steps and tools for applying behavioural science, this chapter — *Test Hypotheses* — is organised slightly differently.

Conducting rigorous evaluations to measure the effectiveness of behavioural interventions is complex and goes beyond the scope of this field guide. Still, this chapter offers essential guidance to help teams understand the value of impact evaluations, the challenges of demonstrating causality, and the key elements and decisions involved in designing and coordinating experiments. The goal of this chapter is to give UNICEF teams/partners a baseline understanding of testing intervention hypotheses to enable more efficient collaboration with evaluation specialists.

Given the wide range of evaluation approaches available, this chapter outlines the most commonly used methods within behavioural science. Many of these activities require advanced technical expertise. It is therefore recommended that teams consult evaluation specialists and use this chapter to understand what expertise may be needed, what to ask evaluation specialists, and which challenges may arise. For teams interested in directly conducting evaluations, additional manuals and practical resources are listed throughout the chapter as well as in the “Learn more” section at the end of this phase.

Why “Test Hypotheses”?

The “Why Test Hypotheses?” section will explore:

1. **The importance of evaluation**
2. **The causation challenge**
3. **The counterfactual framework:** understanding what would have happened otherwise
4. **How randomization creates the gold standard for counterfactuals**

This foundation sets the stage for choosing the right evaluation method and using evidence to strengthen our interventions.

1. The importance of evaluation

Imagine you are ill and a doctor offers you a new medication. When you ask about its effectiveness, the doctor replies: "We haven't tested it formally, but it probably works. Several patients who took it seemed to get better, and our team feels confident about it." Would you take it? Most people would refuse, and for good reason.

Yet when it comes to social programmes and behavioural interventions, we often do the opposite. We implement programmes based on good intentions, promising theoretical frameworks, and anecdotal stories of success, but without rigorous evidence of impact. So why do we hold social interventions to a lower standard than medicine, when both aim to improve human wellbeing?

When success stories mislead

International development has a long history of well-received programmes that captured attention, attracted funding, and seemed to promise breakthrough results — until their impact was rigorously evaluated. One of the most famous examples is microfinance.

Emerging in the 1980s, it was hailed as a transformative tool for poverty reduction. By offering small loans to individuals in low-income settings, microfinance aimed to foster entrepreneurship and economic growth. The model spread rapidly and gained widespread acclaim.

Over time, however, rigorous evaluations told a more complicated story. While microfinance improved access to credit, its impact on poverty reduction, economic mobility, and long-term well-being was less clear. Rigorous research studies highlighted increased debt burdens for borrowers, limited scalability, modest business outcomes, and little progress on addressing structural poverty[1 2 3]. Articles and books — such as "Big Money Backs Tiny Loans That Lead to Debt, Despair and Even Suicide" and More Than Good Intentions reflect the disillusionment that followed and the crucial role evaluation played in revealing what anecdotes could not.

Microfinance is not an isolated case. Other high-profile programmes — like PlayPump[4], One Laptop Per Child[5], and the Millennium Villages project[6] — generated early enthusiasm, but later fell short on their impact when carefully evaluated.

The assumption trap: Why good intentions aren't enough

Even with the best intentions, we're prone to making assumptions that can lead our interventions astray. This "assumption trap" operates at multiple levels:

- **Assuming to understand the problem.** Often, issues are diagnosed from individual perspectives, rather than a deep understanding of the lived experience of communities. What seems obvious from the outside may lack or miss crucial context and complexity.
- **Assuming to know what will work.** Based on personal expertise or past experiences, individuals may become convinced that certain approaches

1 John, B. (2024, November 14). Challenges and limitations of microfinance in achieving large-scale poverty reduction and job creation [Working paper].
2 Akbari, M., Nikijoo, I., Khodapanah, B., Foroudi, P., & Padash, H. (2025). Forty Years of Microfinance Research and Its Impact on Consumers: A Review and Research Agenda Using the ADO-TCM Framework. International Journal of Consumer Studies, 49(4), e70101.
3 Blanc, J. (2014). Microfinance, Debt and Over-Indebtedness: Juggling with Money, Isabelle Guérin, Solène Morvant-Roux et Magdalena Villarreal (dir.). Editions Routledge, Londres, Royaume-Uni, 2014, 316 pages. Revue internationale de l'économie sociale: recma, (334), 122-124.
4 UNICEF. (2007). An Evaluation of the PlayPump® Water System as an Appropriate Technology for Water, Sanitation and Hygiene Programmes https://www-tc.pbs.org/frontlineworld/stories/southernafrica904/flash/pdf/unicef_pp_report.pdf
5 Cristia, Julian and Ibarraran, Pablo and Cueto, Santiago and Santiago, Ana and Severin, Eugenio, Technology and Child Development: Evidence from the One Laptop Per Child Program (February 2012). IDB Working Paper No. IDB-WP-304, Available at SSRN: https://ssrn.com/abstract=2032444.
6 Mitchell, S., Gelman, A., Ross, R., Chen, J., Bari, S., Huynh, U. K., ... Sachs, J. D. (2018). The Millennium Villages Project: a retrospective, observational, endline evaluation. The Lancet Global Health, 6(5), e500–e513. https://doi.org/10.1016/S2214-109X(18)30065-2

will succeed without sufficient evidence. However, it's often possible that the factors that could impact the outcome haven't yet been observed.

- **Assuming implementation will go as planned.** It's common to underestimate practical challenges and overestimate how closely interventions will follow their design when applied in real-world settings.
- **Assuming positive anecdotes mean success.** When favourable feedback is shared or positive moments are observed, it's common to generalize these experiences, giving them more weight than they deserve in assessing overall impact.
- **Assuming correlation means causation.** When things improve after an intervention, it's natural to attribute the change to our work, even when other factors might be responsible.

These assumptions don't stem from carelessness or incompetence, they're a product of how human cognition works. As noted in previous modules, human minds seek patterns, prefer confirming evidence, and create coherent narratives — even when reality is messier. While these tendencies serve people well in many contexts, they can be misleading when evaluating complex social interventions.

Without systematic evaluation, these assumptions remain unchallenged. This may lead to investing in programmes that seem effective but don't actually create meaningful change, or worse, may cause unintended harm. Evaluation provides the structured process needed to move beyond assumptions and understand the true impact of the work.

The value proposition: Why evaluation is worth the investment

Evaluations aren't just academic exercises, they deliver concrete value:

- **Resource optimization:** In resource-constrained environments, evaluation helps direct limited funds toward interventions with proven impact.
- **Course correction:** Timely evaluation allows us to identify and address implementation problems before scaling, preventing the widespread adoption of ineffective approaches.
- **Stakeholder confidence:** Rigorous evaluation builds trust with donors, governments, and communities, facilitating partnerships and long-term support.
- **Scale and replication:** Well-evaluated programmes provide a blueprint for expansion, allowing successful approaches to benefit more communities.
- **Prevention of harm:** Evaluation can identify unintended negative consequences of well-intentioned programmes before they affect large populations.

2. The causation challenge

Beyond "before and after"

When we implement a programme and see improvements, it feels natural to assume our intervention made the difference. A vaccination campaign launches, and disease rates drop; a parent education programme begins, and school attendance rises. These connections seem obvious, but they might be misleading.

The fundamental challenge in evaluation is determining whether our intervention actually *caused* the changes we observe, or whether those changes occurred because of other factors. This is harder than it might appear at first glance.

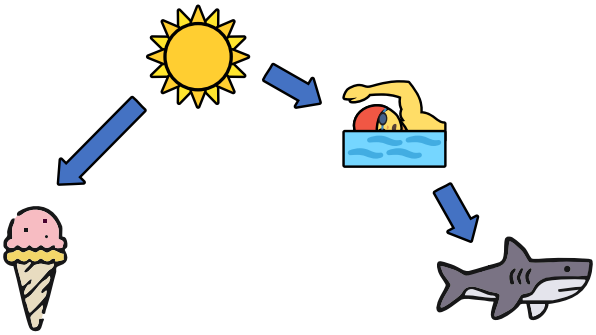
Correlation vs. causation

Correlation means two things happen together. Causation means one thing actually makes the other happen. This distinction is crucial for evaluating the impact of programmes or interventions.

Consider a classic example: ice cream and shark attacks. Data shows that when ice cream sales go up, so do shark attacks. Are ice cream sales causing shark attacks? Of course not. This is what statisticians call the "third variable problem" or "common cause confounding", when a hidden factor influences both variables at the same time. In this case, temperature is the hidden third variable that influences both outcomes independently. During summer months:

- Higher temperatures lead to increased ice cream consumption
- The same higher temperatures lead more people to swim in the ocean
- More swimmers in the water increases the likelihood of shark encounters

This can be illustrated with a simple causal diagram or Directed Acyclic Graph (DAG). The blue arrows represent causal influences. There is no arrow connecting ice cream sales and shark attacks because there's no direct causal relationship between them. They are correlated (they happen together) but not causally linked (one does not cause the other).



Why this matters for programmes: If we don't understand the difference between correlation and causation, we risk drawing the wrong conclusions — and designing or scaling interventions that are not actually responsible for the change we're seeing.

Consider a real-world development example. Imagine that a UNICEF nutrition programme is implemented in several communities. Soon after, children's growth indicators begin to improve. It might seem intuitive to attribute this improvement to the programme. But what else may be going on?

- ➔ Perhaps its harvest season, and food availability has naturally increased.

- Maybe another organization started providing clean water, reducing diarrhoeal disease.
- Or perhaps the government implemented an economic policy that increased household income at the same time.

Any of these factors could explain the improvement -or at least contribute to it. If we assume that our programme caused the change, when it was actually due to other factors, we may invest in or scale up interventions that don't actually work. Worse, we might overlook what really did drive the change, missing opportunities to replicate or strengthen more effective solutions.

Understanding correlation vs. causation helps us avoid these pitfalls. It pushes us to ask better questions, design smarter evaluations, and make more informed decisions.

Confounding factors: Why effects are hard to Isolate

When an intervention is implemented and outcomes are measured, many other factors beyond the programme can influence observed changes. These other unrelated factors are known as confounding variables. By failing to account for these variables, there is a risk of attributing effects to the intervention that were actually caused by something else.

This challenge is called endogeneity: a situation where the relationship between an intervention and its outcome is distorted because other variables are at play. Recognising this helps to understand why simple before and after comparisons can be misleading.

Below are some of the most common confounding factors, illustrated with examples drawn from typical UNICEF programme contexts:

Time-based confounders:

Changes that would have happened regardless of our intervention.

- **Seasonal variations.** Nutrition indicators improve after a feeding programme begins, but the programme launched just before the harvest season, when food is naturally more available.
- **Pre-existing trends.** School enrolment rises after an education campaign, but data shows rates were already increasing steadily due to long-term economic development.

Selection confounders:

Differences between those who participate in a programme and those who don't.

- **Self-selection bias.** Families who join a parenting programme may already be more engaged in their children's development, making the programme appear more effective than it is.
- **Targeting bias.** A WASH programme targets communities with high rates of diarrhoeal disease. Even without the programme, these extreme rates might decline over time simply due to natural variation.

Environmental confounders:

External events or conditions occurring at the same time.

- **Concurrent programmes**. A child protection campaign launches as the government begins stricter enforcement of child labour laws. It becomes difficult to discern which initiative drove the observed changes.
- **Policy changes.** An early childhood nutrition programme rolls out just as a national food subsidy is introduced. Both could be influencing improved nutrition indicators.

Measurement confounders:

Changes in how we track or detect outcomes.

- **Improved monitoring.** After a new reporting system is introduced alongside an anti-trafficking initiative, the number of cases rises. This is not due to a rise in trafficking, but an improvement to the process of detection.

These examples illustrate why it's challenging to determine whether an intervention was the real cause of the observed changes. When multiple factors influence outcomes simultaneously, how can a programme's true impact be isolated?

3. The counterfactual framework: Understanding what would have happened otherwise

At the heart of causal inference lies a seemingly simple question: What would have happened if the intervention had not taken place?

This alternative scenario — where the programme didn't exist — is known as the counterfactual. It represents the benchmark against which the real-world outcome is compared to determine whether the programme truly made a difference.

Consider a child who receives a vaccine and doesn't contract the disease. Did the vaccine prevent illness, or would the child have remained healthy anyway? Both outcomes — seeing the child vaccinated and unvaccinated — can't be observed at once. This dilemma is what scholars call the "fundamental problem of causal inference." It simply isn't possible to observe both the actual and the counterfactual for the same individual.

Instead, the counterfactual is approximated by finding or creating a valid comparison group. This group is as similar as possible to the intervention group and experiences the same external conditions — such as changes in season, economic shifts, or policy reforms — but does not receive the intervention.

If both groups are exposed to the same context, any meaningful difference in outcomes between them can be attributed to the programme itself. This is the foundation of credible evaluation design. A carefully constructed counterfactual helps to move beyond assumptions and confidently answer an important question: Did the intervention make the difference, or would it have happened anyway?

Potential outcomes: A formal way to think about counterfactuals

To help with reasoning around casual impact, statisticians use what's called the potential outcomes framework. This provides a formal structure for thinking about the difference an intervention makes by creating multiple possible realities for each unit. A unit can be a person, household, school, or community, and for each unit, there are two potential outcomes:

- $\mathbf{Y}_1$: The outcome if the unit **receives** the treatment/intervention
- $\mathbf{Y}_0$: The outcome if the unit **does not receive** the treatment/intervention

The causal effect is the difference between these two potential outcomes: $\mathbf{Y}_1 - \mathbf{Y}_0$.

But here is the real challenge; we can only ever observe one of these outcomes for any individual. When a child receives a vaccine, we see what happens with vaccination (Y_1). However, it's impossible to know what would have happened to that same child without the vaccine (Y_0). This unobserved alternative, the counterfactual, remains forever unknown.

This is the fundamental problem of causal inference: the need to know both what happened and what would have happened otherwise, when only one reality can be observed.

Child A (Treatment Group) → Received intervention → Observed outcome: Y_1

Child B (Comparison Group) → Did NOT receive intervention → Observed outcome: Y_0

Causal effect = $Y_1 - Y_0$

The solution to counterfactuals: From individual to group counterfactuals

Since both outcomes cannot be observed for the same person (it's not yet possible to clone people or travel between parallel realities!), the focus can instead be shifted onto groups. By carefully constructing comparison groups that are very similar to one another, it's possible to approximate what would have happened to those in the group that received the intervention and those who did not.

Rather than asking, "Did this specific child gain weight because of the programme?" we instead ask

"On average, how much more weight do children gain when participating in the nutrition programme compared to when they don't participate?"

Now that the focus is on groups rather than individuals, it's time to explore how to create valid group-level counterfactuals. In other words, how to create a comparison group that might be similar across as many observable and unobservable factors/variables as possible.

The challenge of constructing valid counterfactuals

With an understanding of the need to compare groups and not individuals, the next concern is how to create a group that accurately represents what would have happened without the intervention.

Constructing this comparison group (or counterfactual) is one of the most important and challenging steps in causal evaluation. As noted earlier, confounding factors can easily distort conclusions. The quality of the causal inference depends entirely on how well the selected comparison group mirrors the intervention group in all ways except one: they did not receive the treatment or intervention.

This is where evaluation design comes in. Different approaches attempt to approximate what would have happened to our intervention group had they not received the treatment (the counterfactual). Each approach comes with its own trade-offs between rigour, feasibility, and risk of bias. Some designs offer stronger causal claims but require more control or resources; others are more flexible but introduce greater uncertainty.

- **Before-after comparison — Weak counterfactual**

  One common but flawed approach to evaluation is the before-after comparison. This method measures outcomes just before a programme begins and again afterward, attributing any change to the intervention. While simple and intuitive, this approach is highly vulnerable to confounding factors that can influence outcomes over time, independent of the programme itself. These include:

  - **Time-based confounders:** Seasonal variations or long-term trends (e.g., an agricultural training programme shows increased yields, but the evaluation period coincides with the natural growth season).
  - **Environmental confounders:** Simultaneous programmes or policy changes (e.g., a nutrition programme appears successful, but the government simultaneously introduced free school meals in the same area).
  - **Measurement confounders:** The act of measurement influences results (e.g., repeated surveys make households aware of "desired" behaviours like handwashing, prompting changes independent of the programme itself).

  All of these factors can create the illusion of impact, when in reality, the change might have occurred anyway.

  For example, a community health programme launched in April shows improved outcomes by August. The month of August, however, is also the start of the dry season, when waterborne illness naturally declines. Hence, the observed improvement may actually be unrelated to the intervention. The before–after approach essentially assumes that "the same group, earlier in time" can act as its own counterfactual. In dynamic, real-world settings, this assumption rarely holds — making this design a weak basis for drawing causal conclusions.

- **Non-equivalent comparison group — better, but flawed**

  A step up from before-after comparisons is a non-equivalent comparison group — a group that does not receive the intervention, but is observed during the same time period as the intervention group. This approach helps to address many time-based confounders since both groups are exposed to the same external conditions (e.g. seasons, policy shifts, or economic changes).

  However, this design is still vulnerable to selection confounders: differences between the groups that can affect outcomes independently of the intervention. These include:

- **Self-selection:** Individuals who choose to participate may already be more motivated, better resourced, or more health-conscious than those who don't.
- **Administrative selection:** Programmes are often intentionally delivered to areas with the greatest need or the highest potential for success, which can skew comparisons.
- **Baseline differences:** Even before the programme starts, comparison communities may differ in key ways, such as infrastructure, income, or demographics.

Researchers often try to match groups on observable characteristics, but this approach has its limits. Many important factors (e.g. attitudes, aspirations, resilience, or genetics) cannot be observed, but can still drive outcomes. These hidden differences make it difficult to confidently attribute changes to the intervention.

While stronger than before-after design, the non-equivalent comparison approach still falls short of producing high-confidence causal estimates unless additional methods (such as statistical adjustment or natural experiments) are applied carefully.

4. How randomization creates the gold standard for counterfactuals

The need for better counterfactuals.

As noted, both before–after comparisons and non-equivalent group designs are limited. They may help to observe change, but they struggle to isolate what caused that change — especially in the presence of confounding factors, both known and unknown.

Just like the example of ice cream and shark attacks, many real-world relationships are shaped by hidden variables. In programme evaluation, these hidden variables are often numerous, complex, and impossible to fully measure.

As such, it's necessary to find a way to create comparison groups that are balanced on characteristics that are both observed and unobserved.

This is where random sampling comes in. By first randomly sampling different groups (individuals, schools, or communities) from the target population and then randomly assigning which groups receive an intervention, this creates statistically equivalent groups. On average, these groups will be similar across all characteristics, both observed and unobserved, because they have been sampled from the same underlying population in a similar manner. This means that factors such as motivation, baseline health, income, unmeasured beliefs, or community norms will be randomly balanced.

When properly implemented, randomization ensures that the only difference between groups is whether or not they received the intervention. This approach makes it far more likely that any differences in outcomes are due to the programme itself, rather than external or pre-existing differences. For this reason, randomization is referred to as the "gold standard" in causal inference. When done well and with a large enough sample size, this approach provides the strongest possible confidence that the intervention caused the observed change.

Randomization as an approach to create counterfactuals

By randomly assigning which units (i.e. individuals, households, schools, or communities) receive an intervention, the influence of confounding factors is significantly reduced. All types of confounders noted earlier are now evenly distributed across groups by design. In practice, this means:

- Seasonal variations affect both groups equally
- Self-selection bias is eliminated, as participation is assigned, not chosen
- Pre-existing trends unfold similarly across both groups
- Concurrent programmes or policy changes impact both groups at the same time
- Measurement-related effects apply equally across groups

As a result, the only systematic difference between the groups is whether or not they receive the intervention. This allows for attributing any differences in outcomes directly to the intervention itself, rather than any other hidden or external influences.

To provide an example, imagine running a programme to encourage parents to send their children to school. There is a large group of eligible parents, but they are a diverse group, each shaped by different factors:

- Some are wealthy, others face financial hardship
- Some live close to school, others live far away
- Some completed higher education, others have little formal schooling
- Some have flexible jobs, others rigid schedules
- Some strongly value education, others are skeptical
- Some had good schooling experiences, others did not
- Some are highly motivated, others less so

By randomly assigning parents to either receive the intervention or be in the control group (i.e., don't receive the intervention), this ensures that all characteristics are distributed similarly between groups. This includes both observable characteristics, such as income and distance, and unobservable ones, such as beliefs and motivation.

Randomizing ensures that there is an equal distribution of these different traits across the two groups. In this example, any difference in school attendance rates after the intervention can be attributed to the intervention itself, rather than any pre-existing differences between the parents who did and didn't receive it.

The path from good intentions to real impact

Rigorous evaluation isn't just about academic credibility, it's about ensuring that programmes actually improve lives. As noted earlier, well-intentioned interventions can fail to deliver impact, waste precious resources, or even cause unintended harm when relying on assumptions rather than evidence. The difference between correlation and causation matters, as it determines whether solutions that truly work are scaled, or investment is made in programmes that simply happened to coincide with positive change. By constructing valid counterfactuals, ideally through randomization, understanding moves beyond what seems to work towards what actually works, for whom, and why.

This knowledge transforms how to design programmes, allocate resources, and ultimately serve communities. While rigorous evaluation may seem daunting, continuing interventions without knowing their true impact is far riskier.

Common concerns about randomization and responses

While randomization is one of the most rigorous approaches to address the challenge of causality and counterfactuals, it tends to generate concerns. These can be structured around six different categories:

1. Cost and resource

CONCERN: Randomized evaluations are expensive and resource intensive.

RESPONSE:

- While rigorous evaluations do require investment, the cost must be weighed against the value of reliable evidence
- Not all RCTs need to be large scale or expensive — small, focused studies can be cost-effective
- The cost of implementing ineffective programmes at scale is ultimately much higher than evaluation costs
- Existing data sources and clever designs can sometimes reduce costs substantially

2. Time constraints

CONCERN: Randomized evaluations take too long, and UNICEF needs to respond quickly.

RESPONSE:

- Rapid-cycle testing, a structured process of trying out small-scale interventions, measuring results quickly, and iterating based on feedback to refine solutions can provide initial insights in shorter timeframes
- Phased implementation allows for both immediate action and rigorous evaluation
- Time invested in evaluation prevents wasting years on ineffective approaches
- Some outcomes can be measured in the short term (e.g., adherence to medication such as ARVs) while others require tracking over longer periods (e.g., suppressed viral load)

3. Ethical concerns

CONCERN: It's unethical to withhold potentially beneficial programmes from control groups.

RESPONSE:

- When resources are limited, randomization is often the fairest allocation method
- There is no way to know whether programmes work without testing them, some may have no effect or even cause harm
- Phased implementation ensures all participants eventually receive the programme if proven effective
- The ethical imperative to ensure programmes truly help children justifies the process of rigorous testing

4. Political and stakeholder challenges

CONCERN: Government partners or communities may resist randomization.

RESPONSE:

- Framing matters — emphasize the benefits of this type of evaluation, including options for keeping it low-cost or time-bound
- Involve stakeholders early in the design process to address concerns and build ownership
- Create clear procedures for stopping trials if strong evidence of benefit emerges
- Explain that evaluation strengthens advocacy for successful programmes

5. Contextual relevance and generalizability

CONCERN: Results from one context won't apply to others where UNICEF works.

RESPONSE:

- Strategic site selection can improve generalizability
- Measuring implementation factors helps to understand what contextual elements matter
- Even localized evidence is better than no evidence
- Multiple evaluations across contexts can build a broader evidence base

6. Technical capacity

CONCERN: UNICEF staff may lack technical expertise to design and analyse randomized evaluations.

RESPONSE:

- Partnerships with academic institutions can supplement internal capacity
- Investment in staff training builds long-term organizational capability
- Simple randomized designs can be more accessible than complex quasi-experimental methods
- Evaluation specialists within UNICEF can provide technical support across programmes

Design considerations for impact evaluation

Having established why evaluation matters and how causation can be credibly identified, the focus now turns to designing evaluations that ask the right questions, measure the right outcomes, and generate evidence that directly informs decisions. The following section outlines key design considerations for conducting rigorous impact evaluations.

1. Defining research questions

Every evaluation begins with a simple challenge: clarifying the key questions that need to be answered. Begin by gathering key stakeholders and asking:

- What decisions depend on this evaluation?
- What would be done differently if X was known, versus Y?

Focus on 3-5 core questions that directly inform concrete actions. These typically fall into three categories:

- **Effectiveness**: Does the intervention work? How long does it take to work? How soon are effects observed?
- **Mechanism**: How does the intervention work?
- **Targeting**: For whom does it work best?

For each question, specify how the answer will guide decisions. If a vocational training programme increases employment by 10%, will it be scaled? What if it's only 5%? What if it works for men, but not women? Pre-defining these decision thresholds prevents any post-hoc rationalization and ensures that the evaluation generates actionable evidence. Remember to consider both positive and null scenarios – knowing something doesn't work is equally valuable for resource allocation.

Once the key research questions are defined, the next step is to determine how those questions will be answered. This requires translating broad objectives into precise, measurable outcomes and selecting the right indicators that capture real change.

2. Selecting primary outcomes and measurement strategy: Starting with precise questions

The success of the evaluation hinges on measuring the right things in the right way. This starts with transforming broad goals into precisely articulated evaluation questions. A vague question such as "Does this nutrition programme work?" leads to vague answers and unclear decisions. Instead, contrast this with a question such as "Does providing monthly nutritional counselling increase height-for-age z-scores by at least 0.2 standard deviations among children 6-24 months in rural communities within 12 months?"

In other words: "Does [intervention] lead to [specific measurable change, with a threshold if relevant] in [defined population] within [timeframe and context]?" This level of precision guides every subsequent measurement decision and ensures that results are both interpretable and actionable.

Once the evaluation questions have been stated, the next step is to identify the outcome measures that can help to respond to the questions.

3. Choosing what to measure

Choose outcomes that clearly show what the programme is trying to achieve. Focus on measures that directly indicate whether the intervention is making the intended difference. These should be specific enough to measure accurately and important enough to guide programme decisions.

Consider how close the outcomes are to the intervention. Proximal outcomes, those that occur soon after the intervention, are easier to change and measure, but may not capture the ultimate goal. Distal outcomes, those further down the results chain, show real impact, but they usually require larger samples and more time to detect.

For example, a school feeding programme might quickly increase attendance (proximal) but would require much larger data and more time to show improved learning (distal). Both types are useful: short-term outcomes show whether the programme is on track, while long-term ones show whether it's truly making a difference.

Once outcomes are clearly defined, it becomes essential to consider how they will be measured. Different data sources vary in accuracy, cost, and feasibility, and recognizing these trade-offs ensures that measurement decisions strengthen rather than compromise the evaluation's credibility.

Data collection methods and trade-offs

Distinct advantages and limitations of data sources:

Survey data provides flexibility to measure exactly what is needed but also introduces several challenges. Self-reported behaviours often suffer from social desirability bias. For example, parents often overreport vaccinations or underreport harsh discipline. Recall periods matter enormously; asking about events from last week yields different results in accuracy than last year. Survey fatigue can also compromise data quality in lengthy questionnaires. Marginalized groups may be more difficult to reach through phone surveys, and individuals in low-literacy communities may have difficulty understanding certain questions. Ensure cognitive interviews are carried out during piloting to ensure questions are culturally appropriate and understood as intended.

Administrative data, such as school records, clinic registers, or programme databases offers cost-effective, objective measurement but comes with constraints. Efforts are limited to what's already collected, which may not align perfectly with outcomes of interest. Data on certain ethnic groups or marginalized populations may be incomplete or missing. Data quality also varies wildly — some clinics maintain meticulous records while others barely function. Using administrative data often forces the unit of randomization to match the administrative levels of schools or clinics rather than individuals.

Direct observation and behavioural measures provide objective assessment but require careful implementation. Observers need extensive training to ensure consistency. It is important to approach communities with sensitivity, seek consent carefully, and avoid intruding on private or culturally sensitive spaces. Technology increasingly enables unobtrusive measurement (GPS tracking, sensor data) but may not be feasible in all contexts without requiring expensive equipment.

Biomarkers and anthropometric measures offer objectivity for health interventions but require specialized training and equipment. Can cold chains be maintained for blood samples? Will participants consent to invasive procedures? How will measurement errors be handled from different assessors, while ensuring that data collection is respectful and minimally burdensome for participants?

Identifying appropriate data sources is necessary but not sufficient alone; the timing and frequency of measurement are equally critical. Poorly timed data collection can obscure genuine effects or misrepresent programme performance.

4. Timing and frequency of measurement

Outcomes emerge on different timescales. Knowledge might change within weeks, behaviours over months, and health impacts across years. Measuring too early risks finding null effects for interventions that need time to work. Measuring too late risks missing effects that fade or become contaminated by other factors.

Consider multiple measurement points to understand effect dynamics. Does impact grow, plateau, or decay

over time? An initial boost that fades might suggest a need for reinforcement — fade-out effects are common with behaviour change interventions. Gradual change may suggest that effects build up over time and potentially from different processes. Ideally, try to budget for at least one follow-up beyond immediate post-intervention (e.g., six months later) to assess persistence.

Seasonal patterns can confound results if not carefully considered. For example, agricultural outcomes vary by harvest cycles, disease patterns follow seasonal trends, and school-based measures fluctuate with academic calendars. Time measurements to avoid conflating programme effects with seasonal variation, or ensure both treatment and control groups are measured simultaneously. Additionally, try to include data collection across these to understand how the effect interacts with these seasonal variations.

Timing indicates when change occurs, whereas mechanisms explain why. Measuring along the causal chain exposes how impact unfolds, where systems fail, and how programs can be refined.

5. Measuring mechanisms along the causal chain

Understanding why programs work (or don't) is equally as important as knowing whether they work. Mechanism measurement serves multiple purposes:

Theory validation: Do hypothesized pathways actually operate? A handwashing programme assumes: information -> knowledge -> attitude change -> behaviour change -> health improvement. Measuring each step validates or challenges these assumptions.

Failure diagnosis: When outcomes don't improve, mechanisms reveal where chains broke. Did teachers not receive training? Did they receive it but did not understand it? Did they understand but did not implement? Did they implement it but students weren't engaged? Each breakdown point suggests different solutions.

Programme refinement: Rather than abandoning "failed" programmes wholesale, mechanism data enables targeted fixes. If parents received nutrition information but lacked resources to buy diverse foods, adding vouchers might unlock impact.

Generalization: Programmes working through universal mechanisms, such as reminder effects, likely transfer across contexts better than those dependent on specific institutional features.

Don't just measure final outcomes, track intermediate steps. For a nutrition programme aimed at reducing child malnutrition, this could look like: caregiver knowledge (immediate), feeding practices (short-term), child dietary diversity (medium-term), and nutritional status (long-term). Each provides valuable information about how the programme is working.

Understanding mechanisms reveals how change happens, but to interpret those patterns accurately, it's important to ensure that the data truly reflects the populations to serve. Representativeness and inclusion in measurement are essential for generating evidence that captures diverse realities, not just those easiest to reach.

6. Ensuring representative and inclusive measurement

Who is measured matters as much as what is measured. School-based surveys systematically exclude out of school children, who are often the most vulnerable. Phone surveys exclude those without phones, and clinic data misses those not seeking care, or those unable to do so.

Build inclusive measurement strategies from the start, ensuring communities contribute to shaping indicators and data sources. Use multiple data sources to capture different populations. Oversample marginalized groups to ensure adequate representation. Translate instruments into local languages and pilot with diverse respondents. Train enumerators from communities being surveyed to improve rapport and understanding.

Consider whose perspective is being captured. Children, parents, teachers, and health workers may report differently about the same phenomenon. Even within

the same group (e.g., mothers), individuals' experiences may vary depending on their social identity, background, or position within the community. Triangulation across reporters can reveal important dynamics but requires clear protocols for handling discrepancies.

Even the most carefully designed instruments can fall short if they cannot be implemented effectively in routine real-world settings. Ensuring that measurement approaches are feasible, reliable, and field-ready is essential.

7. Practical measurement considerations

During the measurement process, keep these considerations in mind:

Pilot extensively. Never assume that instruments working elsewhere will also function in another context. Pilot with a small group of respondents mimicking actual field conditions. Test skip patterns, timing, translation, and comprehension. Debrief enumerators thoroughly, as they often spot problems that respondents won't mention.

Balance comprehensiveness with feasibility. Long instruments provide rich data but suffer from respondent fatigue, higher costs, and quality degradation. Most impacts can be detected with focused instruments. Reserve lengthy measurement for small-scale mechanism studies rather than large impact evaluations.

Plan for measurement error. All measurement contains error: anthropometric assessment varies between assessors, and test scores depend on testing conditions. Build in quality checks, such as standardization exercises for assessors, repeat measures on subsamples, and validation against external sources where possible.

Document everything. Create detailed protocols specifying exactly how each outcome is measured, coded, and constructed. Future researchers need to understand and potentially replicate chosen measures. Include survey instruments, training materials, and variable construction code in appendices.

Sound measurement practices ensure data quality, but its value depends on the study's ability to detect real effects. Adequate statistical power safeguards against false conclusions, ensuring that data not only describes what was observed, but reveals what truly worked.

8. Determining sample size and statistical power

Statistical power is essentially the evaluation's ability to detect a true effect when it exists, like having a radar sensitive enough to spot an approaching airplane. In contrast, an underpowered study is like searching for something with a dim flashlight, unable to find anything even when it's there. This becomes particularly important when determining how many participants are needed. If the sample is too small, there's a risk of concluding that the programme had no effect when in reality, it did (known as a "false negative"). However, gathering data from more participants than necessary wastes resources. The key factors affecting power include:

- The anticipated size of the programme's effects – bigger effects are easier to detect
- How much natural variation exists in the outcome measure – more variation requires larger samples
- The chosen unit of randomization – randomizing at a community level requires more effort than randomizing individuals

Build in realistic assumptions about attrition (10-20% is common), non-compliance (the treatment group not receiving the intervention), and contamination (the control group accessing the intervention). Each reduces an effective sample size; it's better to recruit 20% more participants than discover the study is underpowered after data collection.

9. Randomization architecture

When testing an intervention, choosing the level at which randomization will take place is critical. This means deciding whether to assign the intervention to individuals, groups, schools, communities, or another unit that makes sense for the project. The chosen level should match how the intervention is delivered in real life. For instance, if a new curriculum can only be introduced to an entire classroom rather than individual students, the classroom becomes the most practical unit for randomization.

The level of randomization also depends on what kind of information can be collected. If data can only be measured at a group level — such as household spending or school attendance — it makes sense to randomize at that same level.

TABLE 8. DIFFERENT TYPES OF UNITS OF RANDOMIZATION

UNIT OF RANDOMIZATION	ADVANTAGES	CONSIDERATIONS	WHEN TO USE	EXAMPLE
Individual-level randomization	Requires smaller sample size compared to other levels; high statistical efficiency.	Risk of spillover effects when individuals interact; logistical challenges in delivering different interventions within the same setting (e.g. community or classroom).	Suitable when interaction between individuals is minimal and the intervention can be easily targeted to specific individuals.	SMS vaccine reminders sent to caregivers randomly selected within a large urban district.
Household-level randomization	Captures household-level decision-making; aligns with how many behaviours and outcomes are shaped.	Spillovers are still possible when neighbours interact; clear definition of "household" is required; analysis may need to account for household size differences.	Appropriate when interventions affect or involve all members of a household (e.g. home visits, conditional cash transfers).	Conditional cash transfers provided to randomly assigned households with children under five.
Community or village-level randomization (cluster randomization)	Reduces risk of contamination or spillovers; often easier to manage politically and operationally.	Requires more clusters (communities) to reach statistical power; high variation between communities increases variance; implementation logistics more complex at scale.	Useful when individuals within communities are likely to influence each other or when interventions are delivered publicly (e.g. community mobilisation).	Community health worker-led immunisation campaigns tested across randomly assigned villages.

UNIT OF RANDOMIZATION	ADVANTAGES	CONSIDERATIONS	WHEN TO USE	EXAMPLE
Facility-level randomization (e.g. schools, clinics)	Practical for institutional delivery settings; aligns with existing organisational structures; suitable for staff-level or facility-wide interventions.	Facilities may vary in size, quality, or staff turnover; catchment area overlap can lead to spillovers; statistical power is limited by the number of facilities available.	Appropriate for evaluating interventions delivered through institutions, especially when individual-level targeting is impractical.	Interpersonal communication training provided to staff in randomly selected health clinics.

10. When randomization isn't possible

Sometimes randomization isn't feasible, due to political, ethical, or practical constraints. Stakeholders may view random assignment as unfair, programs may already be rolled out, or sample sizes may be too small for meaningful randomization. In these cases, quasi-experimental methods attempt to create valid comparisons using statistical techniques to approximate the counterfactual that randomization would have provided.

These approaches work by identifying and controlling for factors that influence both programme participation and outcomes, what is sometimes known as "closing backdoor paths." While these methods can provide valuable evidence, they require stronger assumptions about the data and context. They typically need larger samples, more extensive data collection, and more complex analysis than RCTs. Most importantly, they remain vulnerable to bias from unmeasured factors that randomization would have eliminated.

For more guidance on quasi-experimental approaches, see Appendix 1. Only pursue these methods with expert guidance, as their validity depends critically on context-specific assumptions that are often untestable.

Practical implementation checklist for impact evaluations

Evaluation protocol

Before developing an evaluation protocol, it is helpful to first outline any learning objectives, using a tool like the Learning Agenda. This tool helps define and organize the key questions to be answered in an impact evaluation, and its use is illustrated in Appendix 2 through the **case study on increasing childhood vaccination uptake in Lebanon**.

Once the Learning Agenda is in place, the evaluation protocol can then be built. The evaluation protocol provides a structured guide and plan on how the intervention will be evaluated. Developing an evaluation protocol transforms design decisions into a comprehensive technical document that guides implementation and analysis. Think of it as a contract with a future self, preventing selective analysis and ensuring scientific integrity. A strong protocol pre-specifies every analytical decision before seeing outcomes, protecting against conscious and unconscious bias toward finding positive results.

Essential protocol components include:

- Detailed description of the intervention — what exactly will be delivered, by whom, how often
- Theory of Change with clear causal pathways
- Evaluation questions mapped to specific hypotheses
- Power calculations with all assumptions made explicit
- Precise outcome definitions with exact survey questions (if the outcome measures are survey-based)
- Randomization procedures including stratification variables
- Analytical models with specific regression equations
- Covariate lists determined by theory, not data
- Subgroup analyses with clear rationale
- Procedures for handling attrition and non-compliance
- Robustness checks to test the sensitivity of findings
- Pre-specifying the analysis plan is particularly crucial for the following:
- Primary versus secondary outcomes, to prevent outcome switching
- Subgroup analyses, to avoid mining for significant effects
- Inclusion/exclusion criteria for analysis sample
- Handling of outliers and missing data
- Multiple testing adjustments

Any deviation from the protocol should be clearly labelled as exploratory in reports. For reference, see the case study in Lebanon, where the full evaluation protocol is available.

Consider registering the protocol in public repositories (AEA Social Science Registry, RIDIE, ClinicalTrials.gov) before data collection begins. Registration provides a timestamp, proving pre-specification and enabling the research community to track all studies, not just published successes. Include enough detail that another researcher could replicate the evaluation, while maintaining operational flexibility for necessary field adaptations that don't compromise the core design.

Implementation plan

This is the detailed operational roadmap that translates the evaluation design into real-world action. The plan takes the technical protocol and turns it into day-by-day implementation guidance, specifying who will do what, when, where, and with what resources – throughout the entire evaluation lifecycle. The plan includes the following:

- granular timelines with specific dates and milestones
- clear role assignments with individuals responsible for each task
- resource requirements (staff time, materials, transportation, technology)
- delivery tracking indicators
- contingency plans for common problems
- communication protocols between team members and partners

The Implementation Plan tool can be a simple and helpful template to develop a step-by-step guide on how the intervention will be carried out. An example of how this tool can be used is shown in Appendix 2.

Why it matters

The gap between a brilliant evaluation design and failed execution is usually poor implementation planning. Even rigorous research designs fail when tablets aren't charged, surveys aren't translated properly, or teams don't know who's responsible for participant recruitment. The implementation plan prevents the thousand small failures that can invalidate an otherwise well-designed evaluation. It ensures coordination among multiple actors (research team, implementing partners, government officials, community leaders) who may have different priorities and working styles.

Without clear operational planning, there's a risk of discovering too late that vaccination campaigns conflict with your data collection, key staff are unavailable during crucial periods, or materials weren't printed in time. The plan also provides a management tool for keeping complex operations on track and identifying problems before they cascade into evaluation failure.

Keep these key considerations in mind:

- **Plan around fixed constraints.** Work backwards from hard, fixed deadlines – such as agricultural seasons, school years, budget cycles, or religious holidays – to set realistic timelines, then add a 20-30% buffer time for inevitable delays. If data collection is estimated to take three weeks, plan for four.
- **Assign clear single-point accountability for each activity.** Avoid shared responsibility, which often means no responsibility. When "the team will conduct training" is noted, specify exactly who leads, who assists, and who is accountable if an activity doesn't occur.
- **Track delivery, not just outcomes.** Include simple delivery metrics that can be tracked weekly: the number of participants enrolled versus target, surveys completed per day, intervention sessions delivered as planned. These are different from outcome metrics and focus purely on whether activities occurred on schedule.
- **Create a comprehensive budget.** A detailed budget is one that includes often-forgotten costs like translation services, transport for supervisors doing quality checks, phone credit for follow-ups, refreshments for community meetings, and replacement materials for damaged items.
- **Specify data flow precisely.** Note any data-related considerations, such as how the data will be transferred from paper forms to digital databases, who checks the data and has access, how often, and where it will be stored.
- **Plan for common field problems with specific contingencies**. What if heavy rains prevent travel during the survey period? What if key staff get sick or quit? What if government priorities suddenly change and the partner agency is reassigned? What if participants are busy with harvest when follow-up was initially planned? Use the Implementation Risks and Mitigation tool to systematically document any anticipated risks and the strategies in place to address them. To see this tool in action, refer to Appendix 2 for the case study in Lebanon, which illustrates its practical application.

- **Provide practical tools.** Include templates and standard operating procedures as annexes, so field teams have practical tools, not just abstract plans — this means actual scripts for recruitment, step-by-step guides for data entry, and checklists for intervention delivery.

Securing ethics approval

Ethics approval is the formal process of obtaining institutional review board (IRB) or ethics committee approval, which involves navigating submission requirements, timelines, and institutional procedures. This process is necessary to ensure the evaluation can legally and ethically proceed.

Beyond solely understanding ethical principles, this involves managing the practical bureaucratic process, including:

- identifying which IRB has jurisdiction (a university, ministry of health, UNICEF, or multiple)
- completing required training certifications for all team members
- preparing extensive documentation packages in specific formats
- responding to reviewer comments and requests for clarification
- maintaining compliance throughout the study, including amendments, adverse event reporting, and annual renewals

Why it matters

Ethics approval is legally required for research involving human subjects — proceeding without it can invalidate the entire evaluation, expose institutions to legal liability, and destroy community trust. Many journals won't publish results without proof of ethics approval, and donors increasingly require evidence of ethical clearance before releasing funds. The practical challenge is that ethics review can take 2-6 months, with multiple rounds of revision, and delays here cascade through the entire prospective timeline. A technically perfect evaluation is worthless if data collection cannot begin because ethics approval is pending. Moreover, maintaining ethics compliance throughout implementation requires systems for documenting protocol deviations, reporting adverse events, and ensuring all team members follow approved procedures.

Keep these key considerations in mind:

Start early. Begin the ethics process before finalizing all details — amendments can be submitted later for minor changes. Attaining initial approval starts the clock.

Identify the right authorities. Map out which ethics bodies have jurisdiction. Local university IRBs often require affiliation, national health ministry ethics committees may be needed for health research, multiple approvals may be required for multi-country studies, and some donors have their own ethics requirements.

Allow for realistic timelines. Budget significant time for the process: 2-3 weeks to prepare documents, 4-8 weeks for initial review (longer if a full board review is required), 2-3 weeks for responding to comments, 2-4 weeks for final approval, and potential additional time for local or national approvals.

Prepare a complete submission. Assemble a comprehensive documentation package, including the following: detailed protocol with background, objectives, methods; consent forms in all local languages with back-translations; survey instruments, even if still being refined; CVs and training certificates for all key personnel; data management plans with security measures; compensation structures with justification; risk mitigation and referral procedures.

Understand the level of review likely required, as this affects the timeline. An exempt review (minimal risk, specific categories) takes 2-3 weeks; an expedited review (minimal risk, not exempt) takes 4-6 weeks; a full board review (more than minimal risk, vulnerable populations) can take 2-3 months and only meets monthly.

Avoid common delays. Common reasons for delays include: incomplete applications missing required sections, consent forms using technical jargon or missing required elements, inadequate risk assessment or mitigation plans, compensation that appears coercive, unclear data protection procedures, and missing signatures or institutional approvals.

After approval, maintain compliance. This entails training all new team members on protocol, documenting and reporting any deviations, submitting amendments before

making changes, filing annual continuing review reports, and properly closing the study when it's complete.

Implementation monitoring and process evaluation

This is a comprehensive system for tracking how the intervention is actually delivered in the field, combining regular implementation monitoring routines with systematic process documentation. This involves structured check-ins (daily, weekly, or biweekly depending on intensity) using standardized tools to assess multiple dimensions of delivery, including:

- **Fidelity**: was the intervention delivered as designed?)
- **Reach**: what proportion of the target population received the intervention?
- **Dose**: the frequency and intensity of delivery
- **Quality**: how well was the intervention delivered?
- **Participant engagement**: (did participants understand and act on it?)
- **Contextual factors** affecting implementation

It includes real-time tracking of recruitment rates against targets, monitoring attrition patterns to maintain statistical power, documenting all adaptations made during delivery, and gathering feedback from implementers and participants regarding what's working and what isn't.

Why it matters

Many evaluations find no impact – not because the intervention doesn't work, but because it was never properly delivered. A lack of systematic monitoring poses a risk: for instance, after expensive endline data collection, it could be discovered that half of the treatment group never received the intervention, control groups accessed the treatment, or field staff modified the intervention beyond recognition. Process evaluation distinguishes between **theory failure** (the intervention genuinely doesn't work, even when delivered well) and **implementation failure** (it wasn't delivered properly, therefore its effectiveness cannot be judged).

This information is crucial for interpreting results. For example,if no impact was found, was it because the theory was wrong or because only 30% of participants received the full intervention? If positive effects were found, understanding what was actually delivered aids in replication efforts. Real-time monitoring allows for course correction while there's still time. For example, if recruitment is falling behind, efforts can be intensified before it threatens statistical power; if certain sites aren't delivering properly, additional support can be provided; if unexpected barriers emerge, solutions can be developed.

Process data also reveals critical insights for scaling decisions: which contexts facilitated smooth delivery, what adaptations were necessary, which implementation challenges are likely to persist at scale, and what level of quality is realistically achievable in routine conditions versus research settings.

Keep these key considerations in mind:

- **Establish monitoring routines that match implementation intensity but don't overwhelm field teams.** This can entail daily huddles for intensive interventions, weekly calls for standard programs, or monthly reviews for light-touch interventions.
- **Create simple tools**. Standardized monitoring sheets can capture essential information without creating an excessive burden of paperwork.
- **Track core implementation metrics separately from outcome data**. This can entail the per cent of planned sessions delivered, the average attendance or participation rates, per cent receiving full intervention dose, time between intervention components, or quality ratings from standardized observations.
- **Set up systems for data quality monitoring**. These can include high-frequency checks (automated daily or weekly data review for outliers, missing data, and suspicious patterns) along with back-checks (re-interviewing 10-20% of participants to verify data accuracy and catch any potential fraud).
- **Monitor assumptions affecting statistical power continuously**: Consider recruitment rates versus targets (is the project on track to reach the sample size?), overall attrition rates (what per cent was lost to follow-up?), differential attrition (is dropout higher in treatment or control?), compliance rates

(what per cent of the treatment group is actually receiving intervention?), and contamination (is the control group accessing the intervention?).

- **Document every adaptation using structured frameworks.** Specifically note what changed from protocol, why the change was necessary (barrier encountered, stakeholder request, feasibility issue), when did the change occur, who made the decision, was the change planned or reactive, and finally, if this affects the core intervention theory.
- **Create rapid feedback loops with implementers.** Use simple WhatsApp groups for real-time problem-solving, brief weekly check-in calls focusing on challenges and solutions, and monthly reviews of monitoring data to identify patterns.
- **Distinguish between core and adaptable components**. Differentiate between core components that must be maintained for the intervention theory to hold versus peripheral elements that can be adapted to context. Document both types, but treat them differently in analysis.
- **Gather participant feedback.** Build in participant feedback mechanisms through brief exit interviews after intervention sessions, periodic focus groups with participants, and suggestion boxes or hotlines for ongoing input.
- **Track contextual factors that might affect implementation or outcomes.** Note other concurrent programs or policies affecting the target population, seasonal factors (holidays, agricultural seasons, and weather), political or security situations, and health emergencies or other disruptions.
- **Maintain detailed logs.** These become crucial for interpreting results, informing scale-up decisions, and contributing to the broader evidence base on implementation challenges and solutions.

Cost-benefit analysis

This is a systematic calculation of all resources required to deliver the intervention and achieve measured impacts, producing standardized metrics that enable comparison across different interventions, delivery models, or investment options. This comprehensive accounting goes beyond simple programme budgets to capture the full economic cost of achieving outcomes, including:

- direct programme costs (staff, materials, operations)
- indirect costs often hidden in other budgets (supervision, administration, overhead)
- opportunity costs of resources used (volunteer time, government staff, participant time)
- startup versus running costs
- marginal costs of adding participants

The analysis produces metrics such as cost per child vaccinated, cost per percentage point increase in test scores, cost per year of life saved, or return on investment ratios that decision-makers can compare against benchmarks or alternative interventions. See the Cost-Benefit Analysis tool here to help identify, quantify, and compare intervention costs and benefits.

To see an example of how to apply this tool using the case study in Lebanon, refer to Appendix 2.

Why it matters

Even highly effective interventions may not be worth scaling if they're prohibitively expensive relative to alternatives. Cost-benefit analysis transforms evaluation from an academic exercise proving something works into actionable policy guidance, answering "Is this good value for money?" This is increasingly critical as donors demand evidence of both impact and cost-effectiveness, governments with limited budgets need to maximize outcomes per dollar spent, and scaling decisions require understanding how costs change with scale.

A programme achieving 10% improvement might seem successful, until realizing it costs five times more than an alternative achieving 8% improvement. Understanding cost structures also reveals opportunities for efficiency — perhaps 80% of impact could be achieved at 50% of cost by streamlining certain components, or fixed costs that seem high for a pilot would be negligible at scale. Without rigorous costing, programs may be abandoned as "too expensive" based on incomplete understanding, or scaled enthusiastically without recognizing unsustainable cost structures.

Keep these key considerations in mind:

- **Start early**. It's important to collect cost data from day one of implementation — retrospective cost reconstruction is unreliable and often impossible as receipts are lost, staff forget time allocations, and in-kind contributions go undocumented.
- **Capture all costs, not just those in the budget.** These costs can include staff time including preparation and training (even if paid by partners), volunteer time valued at local wage rates for equivalent work, government staff time even if not paid by project, participant costs (transportation, lost wages, childcare), donated materials or venues at market value, and organizational overhead reasonably attributed to the project.
- **Distinguish between different cost categories, as they behave differently at scale**. These include fixed costs (training development, initial setup) that don't increase with participants, variable costs (per participant materials, incentives) that scale linearly, and step costs (supervision, new sites) that jump at thresholds.
- **Track costs from multiple perspectives, as different stakeholders care about different numbers**. Consider the implementer perspective (what does it cost this team to run?), government perspective (what would it cost to integrate into existing systems?), societal perspective (including all costs, regardless of who pays), and participant perspective (what does participation cost beneficiaries?).
- **Calculate multiple cost-effectiveness metrics to enable different comparisons**. This can include the cost per participant reached/enrolled/completing; cost per unit of primary outcome achieved; cost per standardized effect size for academic comparisons; incremental cost-effectiveness ratios if comparing variants.
- **Compare results to relevant benchmarks**. Take note of similar interventions in the chosen context, the government's revealed willingness to pay for similar outcomes, international standards (such as WHO thresholds for health interventions), and alternative ways to achieve similar goals.
- **Include sensitivity analyses showing how cost-effectiveness changes under different assumptions**. Consider if effects persist for two years versus one year, if the intervention is implemented by government versus NGO salaries, at different scales (100, 1,000, 10,000 participants), or with different compliance or attrition rates.
- **Document cost drivers transparently. What makes this intervention expensive or cheap?** Could specific components be modified to reduce costs without sacrificing effectiveness? What economies or diseconomies of scale are likely? What hidden costs might emerge in routine implementation versus research conditions?
- **Present findings in accessible ways for policy audiences**. This can include simple cost per outcome metrics, not complex economic models, visual comparisons to familiar interventions, clear statements about confidence intervals, and practical implications for budget planning.
- **Interpret cost-effectiveness wisely.** Remember that cheapest isn't always best — sometimes higher-cost interventions are worthwhile if impacts are proportionally larger or if they reach populations that cheaper alternatives miss.

Rigorous testing strengthens the link between design and decision-making. By grounding conclusions in evidence rather than assumptions, impact evaluations enable UNICEF and partners to start identifying which interventions are effective, which require adaptation, and which should be discontinued. Through systematic measurement and transparent analysis, behavioural insights and design hypotheses are translated into credible, actionable evidence that informs decisions about scale, policy, and resource allocation.

As programmes move into the Scale phase of the DEPTHS process, generated evidence from Test Hypotheses is applied. The findings from rigorous testing guide how interventions are refined, integrated into systems, and expanded responsibly. This ensures that intervention scale-up and replication decisions are grounded in demonstrated impact, rather than assumption.

Appendix

APPENDIX 1:

Quasi-experimental alternatives when randomization is not feasible

This appendix provides additional technical guidance for practitioners designing impact evaluations under real-world, routine conditions. While the main chapter highlights randomized designs as the most reliable method for establishing causality, practical realities sometimes limit their use. In such cases, evaluators may need to consider alternative approaches to randomization that still aim to generate credible, evidence-based insights.

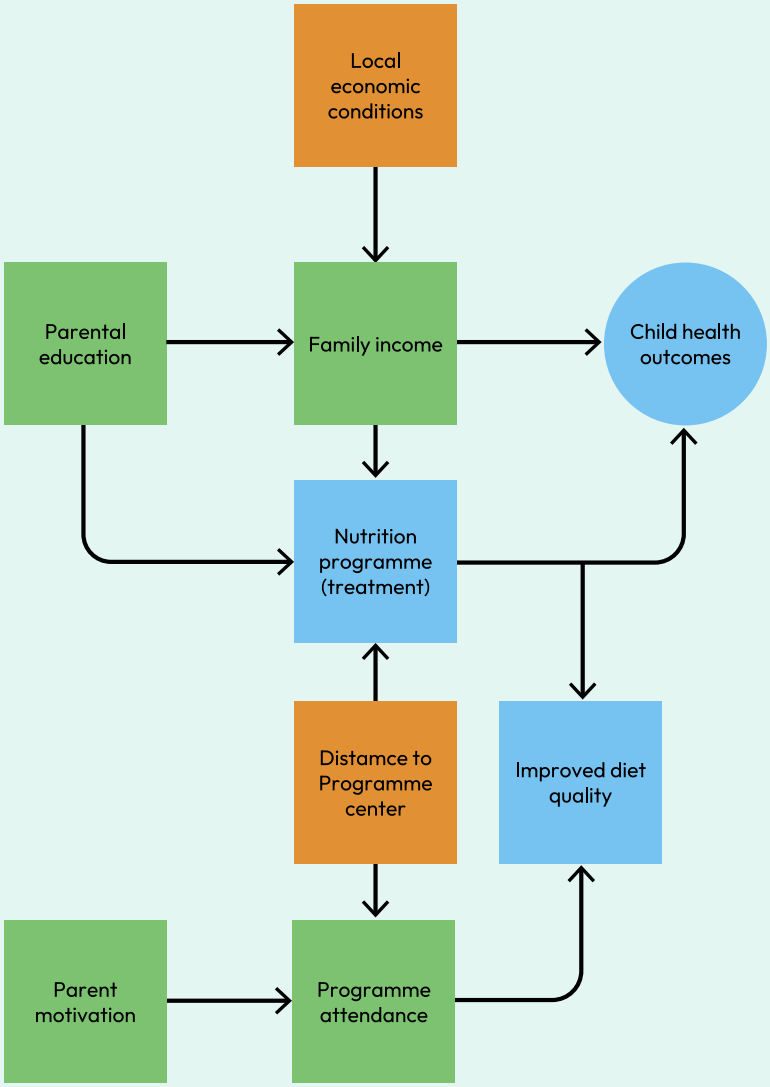
The priority in evaluation design should be to randomize, as it remains the most reliable method for establishing causality. Randomization eliminates systematic differences between groups, allowing for greater confidence that observed effects are attributable to the intervention itself. However, there are circumstances in which randomization is not possible due to political, ethical, logistical, or practical constraints. In such cases, alternative approaches may be considered.

While these alternatives can still yield valuable insights, they introduce both greater operational complexity and increased statistical demands. More importantly, such methods also carry a heightened risk of bias. Engaging experts in causal inference and evaluation design is strongly recommended when non-randomized designs are pursued.

1. **Closing backdoor paths: A DAG perspective**. When randomization is not possible, it becomes essential to identify and control for confounding factors — an approach known in causal inference as "closing backdoor paths." **Directed acyclic graphs** (DAGs) provide a useful framework for understanding this process.

   Consider a DAG illustrating the evaluation of a community nutrition programme targeting children (right). In such a diagram:

   - Arrows represent direct causal effects
   - Variables (nodes) denote factors that influence outcomes or programme participation
   - Paths between variables represent potential causal or non-causal associations



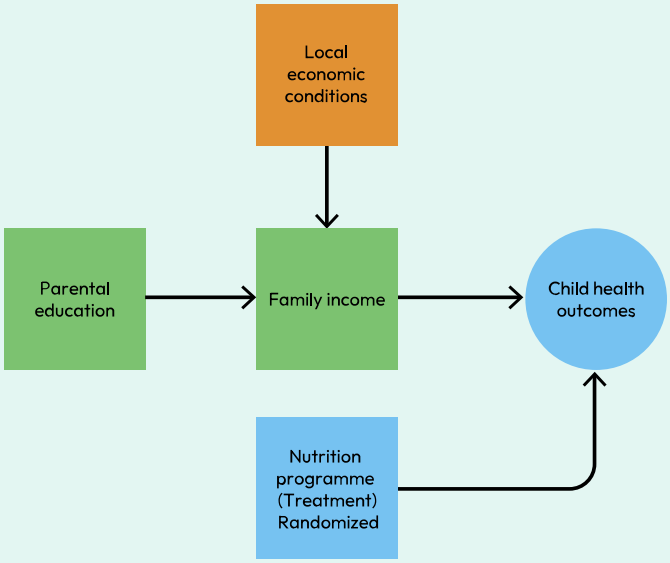
2. **Identifying backdoor paths.** A "backdoor path" refers to any pathway between the treatment (nutrition programme participation) and the outcome (child health outcomes) that does not follow the direct causal direction. These paths introduce associations that can bias causal estimates.

   In the example DAG, several backdoor paths may exist:

   - **Programme → Family Income → Child Health:** Families with higher income may be both more likely to participate in the programme and to have healthier children, regardless of the programme.
   - **Programme ← Parental Education → Child Health:** More educated parents may enrol more frequently and also provide better care.
   - **Programme ← Parental Education → Family Income → Child Health**: Parental education influences income, which in turn affects both participation and health outcomes.

3. **The goal: Closing all backdoor paths.** To isolate the causal effect of the nutrition programme, all backdoor paths must be closed. A path is considered closed when:
   - **A variable on the path is controlled for (conditioned on)**. For example: Controlling for family income closes the path Programme → Family Income → Child Health.
   - **The path contains a collider**. A collider is a variable influenced by two or more variables. For instance, "Programme Attendance" may be influenced by both "Distance to Programme Centre" and "Parent Motivation." This path is naturally closed unless the collider is incorrectly conditioned on, which would re-open the path.
   - **The path includes a mediator that is intentionally left uncontrolled**. For example, "Improved Diet Quality" lies on the causal chain between programme participation and child health. If the total effect is of interest, this variable should not be controlled for.

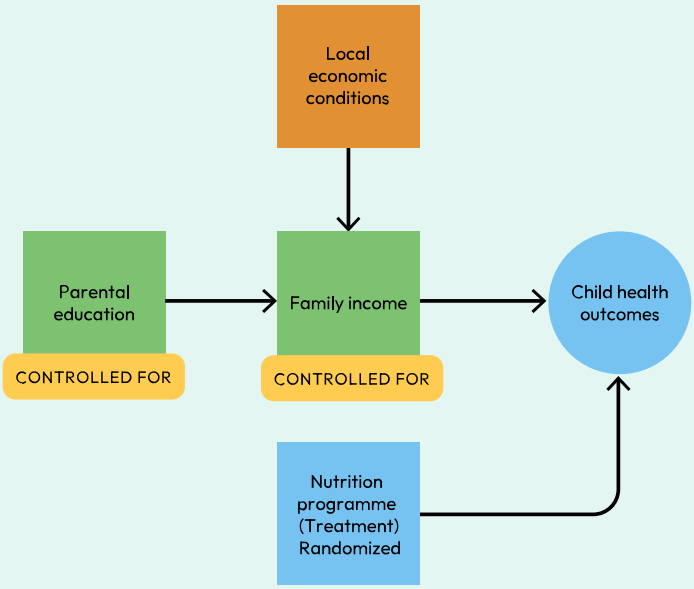
   When randomization is applied, all arrows pointing into the "Programme Participation" node are effectively cut. This renders participation independent of all confounders and closes all backdoor paths at once. This is the core advantage of randomization — it eliminates the need to identify or measure every potential source of bias.



4. **Closing backdoor paths without randomization.** In the absence of randomization, statistical methods can close backdoor paths. Two common strategies include:

   a. **Controlling for observed confounders.** Measure and adjust for variables such as parental education and family income through methods like regression adjustment or matching. This approach requires:

      - Identification of all relevant confounders
      - Accurate measurement of those confounders
      - Correct modelling of their relationships with treatment and outcomes



   b. **Using instrumental variables.** This method relies on a variable that influences programme participation but is not directly related to the outcome. For example, "Distance to Programme Centre" may determine participation without affecting child health directly. This variation can be used to estimate causal effects even in the presence of unobserved confounding.

   The opportunities and limitations of each method depend on the context and the availability of data. Different situations will offer different leverage points for closing backdoor paths, and choosing the appropriate approach requires careful consideration of both design and data constraints.

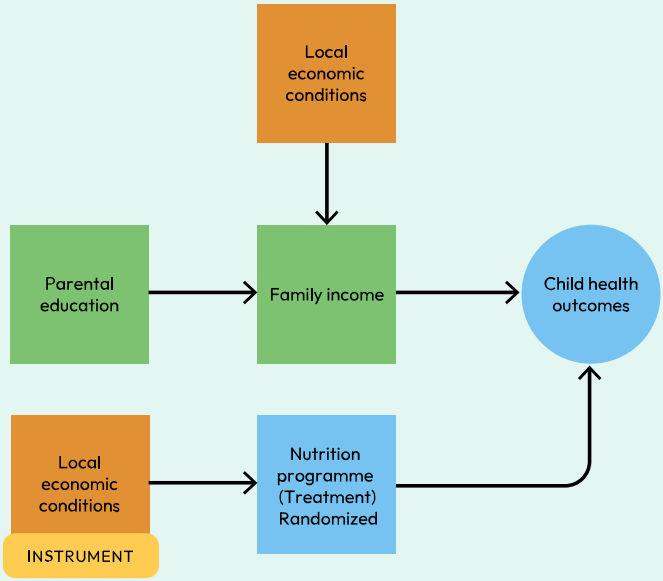


Table 3 in Step 1 of this chapter includes a description of different approaches when randomization is not possible.

APPENDIX 2:

The case of increasing childhood vaccination uptake in Lebanon

This phase also uses the case study introduced in previous phases, aimed at increasing childhood vaccination in low resource settings in Lebanon, in order to illustrate tools in the DEPTHS process.

Context of the case study:

Through the Define, Explore & Diagnose, and Prototype Designs phases, the Lebanon team identified a target behaviour to change (caregivers returning with their child for the next scheduled vaccination) and its main behavioural barriers (cognitive overload from competing tasks, emotional stress related to clinic visits, and social perceptions shaped by community views). They also tested a prototype of a potential solution: a paper-based appointment card intended to make the next vaccination visit salient and easier to remember.

The intervention was tested through a randomized controlled trial involving 12,332 un- or under-vaccinated children across 6,160 households. Households were randomly assigned to either a treatment group that received the appointment card during outreach visits or a control group that did not. The primary outcome was defined at the household level: whether at least one eligible child in the household received the next required vaccine within 21 days of becoming due.

This binary, household-level measure was selected to better reflect caregiver decision-making, recognizing that the marginal cost of vaccinating additional children in the same household is typically low. The analysis used logistic regression to estimate the intention-to-treat effect, with robust standard errors clustered at the outreach worker level. Baseline covariates included household size, nationality, and previous immunisation behaviour.

The trial found that households receiving the appointment card were significantly more likely to return for vaccination, with an increase of 6.7 percentage points compared to the control group. These results demonstrate that a behaviourally informed, low-cost intervention can measurably improve childhood immunisation uptake when targeted at key decision points. The findings have informed ongoing policy discussions and demonstrate the value of structured hypothesis testing in behavioural public health programming.

Per the field guide narrative, this appendix includes a completed example to illustrate its practical application and to support teams in using the referenced tools with ease and consistency. These tools include:

- Learning Agenda
- Implementation Plan
- Implementation Risk Mitigation
- Cost Benefit Analysis

Application of the Learning Agenda

This Learning Agenda was not developed by the original project team. It is a recreated example based on real project data and context.

The following illustrates how to apply the **Learning Agenda** introduced in the field guide, using the case study in Lebanon. This tool offers how to frame and prioritize key learning questions that guide both evaluation design and interpretation of results. The example highlights how a clear Learning Agenda — focused on what the team needs to learn, why it matters, and how findings will be used — ensures that evaluation is purposeful, actionable, and adaptive.

Learning Agenda

D E P TEST HYPOTHESES S

DEPTHS TOOLKIT

Learning Agenda

Use this worksheet to define the key elements of your evaluation plan using the PICOS framework.

Intervention: _____

Problem *What is the problem that the intervention will address?*

PICOS

Population *Who will the intervention be addressed at?*

Intervention *What is the specific intervention that will be implemented and evaluated?*

Comparison *What is the counterfactual group?*

[ ] Randomized Control Group

[ ] Non-Random Control Group (Quasi-experimental methods)

[ ] No Counterfactual (i.e., no control group)

Outcome *What are the main outcome measures that will be used to know if the intervention has been effective?*

Study and evaluation type *What type of evaluation approach will be used and what type of study?*

[ ] Impact

[ ] Process

Key lessons

How will the results be interpreted and used to improve the intervention and the outcome statement?

If the results are positive:

If the results are null or it is inconclusive:

If the results are negative (the intervention backfires):

The Lebanon team first developed a simple Learning Agenda to clarify what they hoped to learn in the evaluation and why. They identified a specific problem: caregivers in Lebanese and Syrian communities were missing routine vaccinations for their children, often because they forgot or misunderstood when to return. The intervention, a simple, behaviourally designed postcard, was meant to address this by serving as a physical reminder.

The primary learning question they posed was: "Does providing a personalized appointment reminder card increase timely childhood vaccination uptake among caregivers in low-income Lebanese and Syrian communities?"

Secondary questions included how the intervention influenced caregivers' intention to return and their recall of the follow-up date. Anticipated interpretations of results were also mapped.

For example, if uptake increased, the intervention could be scaled. If results were mixed or null, design tweaks and further testing might be needed. If backfiring occurred, the team would investigate unintended effects like mistrust or misinterpretation.

Application of the Implementation Plan

This Implementation Plan was not developed by the original project team. It is a recreated example based on real project data and context.

The following illustrates how to apply the Implementation Plan introduced in the narrative of the field guide, using the case study in Lebanon. This tool demonstrates how to translate evaluation design into structured operational planning — defining key activities, roles, timelines, and monitoring indicators. The example provides a practical reference for developing implementation plans that promote coordination, transparency, and accountability throughout the evaluation process.

Developing an Implementation Plan

To operationalize the evaluation, the Lebanon team also developed a clear Implementation Plan outlining the practical steps, responsibilities, and timeline for rollout. This plan detailed each phase, from finalizing the postcard design and conducting a small pre-test, to collecting baseline data and assigning participants to treatment or control groups. Each stage was mapped with corresponding leads (e.g. the UNICEF project team, J-PAL MENA, the Ministry of Public Health), a defined timeline, and specific indicators for tracking progress, such as enrolment rates, distribution coverage, and data quality checks. This plan served not only as a coordination tool but also as a foundation for transparency and accountability throughout implementation.

Equipped with the Evaluation Plan and Implementation Plan, the team was ready to submit to an Institutional Review Board for an ethics review and approval before they implemented and evaluated the intervention.

D E P TEST HYPOTHESES S

DEPTHS TOOLKIT

Implementation Plan

Intervention: Appointment reminder card

Use this worksheet to break down your intervention into actionable steps. For each priority area, define the key activities, assign responsibilities, set a timeline, and identify what resources and monitoring indicators will be needed to track progress.

Priority area *[Insert Priority]*	**Key activities** *[List 1–2 actions]*	**Responsible** *[Lead & supporters]*	**Timeline** *[MM/YY]*	**Resources and budget** *[Inputs needed]*	**Monitoring indicators** *[e.g., % of activities done]*
Pilot conducted	Finalise postcard design and conduct small pretest in 1–2 clinics	AIA programme team, MoPH design unit, UNICEF project team	03/23	Design consultant, transport	Postcard design finalised and feedback integrated
Baseline data collection	Collect baseline data on caregiver return rates and clinic performance	Evaluation team (J-PAL MENA), MoPH, UNICEF project team	03–04/23	Enumerator time, travel, tablets	% baseline surveys completed, data quality checks
Recruitment conducted	Identify and enrol eligible caregivers during routine clinic visits	Health workers, clinic admin staff, UNICEF project team	05–07/23	Staff orientation, clinic rosters	# of caregivers enrolled; consent rate
Assignment conducted	Randomly assign caregivers to postcard (treatment) or standard care (control)	Evaluation team, AIA implementation support, UNICEF project team	05–07/23	Randomisation protocol, data forms	Randomisation completed, balance checks verified
Intervention begins	Begin distribution of postcards to treatment group caregivers post-vaccination	Clinic staff, supervised by AIA field team, UNICEF project team	05/23	Printed postcards, tracking forms	% of eligible caregivers receiving postcard
Implementation check	Conduct midline supervision calls and field visits to assess fidelity and reach	AIA field team, MoPH district focal points, UNICEF project team	06/23	Phone/data, site visit costs	% of clinics reporting on routine, deviations noted
Intervention ends	Conclude distribution phase and stop enrolment of new participants	Clinic staff, evaluation team notified, UNICEF project team	07/23	N/A	Cut-off date enforced across all pilot sites
Data collected	Administer follow-up caregiver surveys; extract clinic return records	J-PAL enumerators, clinic M&E staff, UNICEF project team	08–09/23	Survey tools, transportation, incentives	% follow-up surveys completed, records extracted
Data cleaning and analysis	Clean datasets, conduct statistical analysis on primary and secondary outcomes	Evaluation analysts (J-PAL MENA), UNICEF project team	10–11/23	Analyst time, software	Final analysis plan completed; findings validated

Application of the Risk Mitigation Tool

The following illustrates how to apply the **Implementation Risks and Mitigation Tool** introduced in the main field guide, using the case study in Lebanon. This tool demonstrates how potential operational, contextual, and behavioural risks can be systematically identified and managed throughout the evaluation process. The example provides a practical reference for teams seeking to anticipate and address implementation challenges before they threaten evaluation integrity or programme success.

Navigating risks during rollout

The team used the Implementation Risks and Mitigation tool dynamically, not only at the planning stage but as part of active implementation. Five key risks emerged:

RISK	LIKELIHOOD	IMPACT	MITIGATION
Postcards not arriving on time	Medium	High	The UNICEF team worked with a courier service to stagger delivery schedules and provide buffer stock.
Nurses not delivering postcards with the right explanation	High	Medium	Voice note training was implemented and laminated cheat sheets were distributed.
Caregivers misplacing the postcard	Medium	Medium	The use of clear plastic sleeves and reminders to stick the postcard near the calendar at home.
Data logs not being consistently updated	Medium	High	A quick log reminder was integrated at the end of each shift via WhatsApp prompts.
Contextual disruptions (e.g. transport strike, weather)	Low to Medium	High	Local focal points were assigned for flexible reallocation of delivery routes.

Each risk was assigned to either the UNICEF team or a MoPH supervisor for follow-up, with deadlines linked to the implementation timeline.

Application of the Cost-Benefit Analysis

The following illustrates how to apply the Cost-Benefit Analysis Tool introduced in the narrative of the field guide, using the case study in Lebanon. This tool demonstrates how programme costs and benefits can be systematically identified, quantified, and compared to assess value for money. The example provides a practical reference for teams seeking to estimate the economic efficiency of behavioural interventions and to inform decisions about scaling, adaptation, or resource allocation.

Cost-Benefit Analysis

The primary goal of the intervention was to increase timely vaccination coverage among caregivers in low-resource urban clinics in Lebanon. In the absence of the intervention, data suggested that many caregivers would delay return visits, risking missed or incomplete immunization. Offering a simple behavioural prompt, the postcard aimed to change that trajectory.

Costs were relatively modest. While the study did not report financial data directly, a reasonable estimate placed the cost per printed postcard at under $0.20, including design

D E P TEST HYPOTHESES S

DEPTHS TOOLKIT

Cost-Benefit Analysis

Intervention: Appointment reminder card

Use this worksheet to weigh the costs and benefits of your intervention and decide whether it's worth scaling, adapting, or stopping.

Define the Purpose

The project aimed to improve on-time childhood vaccination through a simple, low-cost reminder postcard. Caregivers—especially low-literacy and refugee families—benefited by receiving a clear next-visit reminder. Without the intervention, drop-out rates and missed vaccinations remained high.

List All Costs

- Direct: Card design, printing, training, and distribution (~$0.20/card).
- Indirect: Staff time, supervision, and monitoring.
- Opportunity: Minimal, as delivery occurred during existing visits.
- Total: Estimated <$10,000 for pilot phase.

List All Benefits

- Direct: More children vaccinated on time.
- Indirect: Reduced dropouts, fewer missed appointments, better caregiver engagement.
- Equity: Stronger impact among Syrian and low-literacy groups.

Assign Values

- Costs valued via procurement/staffing rates.
- Benefits: 7 percentage point increase in timely returns, ~350 additional children vaccinated.
- Proxy value per timely vaccination: ~$50–$150 (based on WHO estimates).

Compare Costs and Benefits

- Cost per additional timely vaccination: ~$28
- Estimated BCR: ~3:1
- Strong return on investment.

Test Assumptions

Sensitivity checks showed results held under:

- Lower effect sizes (3–5%)
- Higher postcard costs
- Even in worst-case scenarios, benefits exceeded costs.

Make a Judgment

The intervention is cost-effective, scalable, and equitable. Recommended for scale-up in similar low-resource settings, with continued feedback loops to optimise delivery.

and distribution. Assuming a modest pilot scope (e.g. <10,000 caregivers), the total cost likely remained under $10,000, including materials, supervision, and staff time.

On the benefits side, the study demonstrated that the postcard increased return rates by seven percentage points. Public health literature suggests that each additional timely vaccination contributes to long-term benefits in disease prevention, reduced child mortality, and lower healthcare costs. Using WHO estimates, a conservative valuation of each timely vaccination could be $50–$150 in societal benefits. When applied across the intervention group, this translates into significant aggregate gains, likely yielding a benefit-cost ratio between 3:1 and 5:1.

To test robustness, the team considered a range of assumptions. Even under pessimistic scenarios (e.g. higher costs or weaker effects), the intervention still appeared cost-effective, largely due to its low unit cost and scalable design.

Importantly, the team also considered equity and inclusion. The intervention disproportionately benefited marginalised caregivers, especially Syrian families, highlighting its potential as a low-cost strategy for reducing health access disparities. The reminder postcard was not only cost-effective but also equity-enhancing, a key consideration for future scale-up.

Learn more

This field guide offers practical tools, frameworks, and worksheets to help teams apply behavioural science to real-world development challenges. However, no guide can cover everything. Behavioural science sits at the intersection of multiple disciplines, ranging from human-centred design and implementation science, to ethics, measurement, and evaluation. That's why we've included this section – for those who are curious to dig deeper, sharpen their ethical practice, strengthen implementation design, or explore how to select better outcome measures. The resources below offer curated starting points for a self-paced learning journey.

"I want more detailed step-by-step guidance on how to conduct experiments."

There are multiple manuals, resources, and courses on conducting experiments for social programmes. Some helpful free resources are the J-PAL's website on Introduction to randomized evaluations and the World Bank's Impact Evaluation in Practice Guide.

"I want to improve how I approach ethics in applied behavioural science."

Ethics is foundational to any research or behavioural project involving people. Whether the task at hand is writing consent forms, evaluating risk, or navigating power dynamics, these resources provide accessible and practical support:

- UNICEF's Ethics Toolkit for Applied Behavioural Science Projects helps teams to reflect on ethical risks early and integrate safeguards throughout implementation.
- UNICEF's Procedure on Ethical Standards in Research, Evaluation, Data Collection, and Analysis outlines the organization's official protocols and expectations.
- Informed Consent Checklist (J-PAL) is an annotated template with guidance on what to include in participant consent forms.
- UNICEF Consent Templates (see page 41) include editable samples for participants, caregivers, and gatekeepers.

"I need to take an ethics course for IRB."

There are multiple training sessions available, with some organizations even offering their own internal ethics training with a certificate allowed by different IRBs. For those seeking external training, explore the following resources:

- HHS Human Research Protection Training (US-based, free certification, ~5–6 hours)
- Tri-Council Policy Course on Research Ethics (Canada-based, free certification, ~4 hours)

"I want to explore how to design and measure implementation more effectively."

Understanding what exactly was accomplished, along with how, is essential to recognizing whether a behavioural intervention worked. The Implementation Outcome Repository offers guidance and examples for measuring constructs like feasibility, fidelity, and acceptability.

"I want to improve the way I select or adapt outcome measures."

A strong outcome measure doesn't just test effectiveness — it captures the right behaviour in the right way. If planning to test behavioural change or proxy outcomes, the Psychometric Properties of Implementation Measures reviews validity and reliability of commonly used tools in implementation science.

"I want to assess the quality and rigour of evaluation reports and studies."

If responsible for reviewing, commissioning, or interpreting studies, it's important to understand not only what a report says, but how trustworthy its findings are. These tools and articles help to assess study quality, whether reviewing an impact evaluation, implementation report, or academic article.

Assessing overall design and reporting rigour

- Gates' DAC Assessment Tool (DAT) helps to assess whether a study was well-designed, well-analysed, and clearly communicated. Originally built for clinical trials, it's applicable across sectors.
- Publishing Quantitative Papers with Rigor and Transparency article offers digestible guidance on transparency and robustness for teams writing up results.

Reviewing systematic reviews

The Evidence Project Risk of Bias Tool evaluates rigor in both randomized and non-randomized studies. It's particularly useful when reading systematic reviews or mixed-methods syntheses.

Evaluating qualitative research rigor

- Indicators of Rigor in Qualitative Research explains how to judge the credibility, transferability, and dependability of qualitative studies.
- Information Power in Qualitative Sampling offers a helpful alternative to the idea of "saturation" for justifying sample sizes in interviews.

Resources

1. Akbari, M., Nikijoo, I., Khodapanah, B., Foroudi, P., & Padash, H. (2025). Forty Years of Microfinance Research and Its Impact on Consumers: A Review and Research Agenda Using the ADO-TCM Framework. International Journal of Consumer Studies, 49(4), e70101.
2. Behavioural Insights Team. How to Run Simple Behavioural Insight Projects. 2022. https://www.bi.team/wp-content/uploads/2022/11/BIT-Handbook-How-to-run-simple-BI-projects.pdf.
3. Blanc, J. (2014). Microfinance, Debt and Over-Indebtedness: Juggling with Money, Isabelle Guérin, Solène Morvant-Roux et Magdalena Villarreal (dir.). Editions Routledge, Londres, Royaume-Uni, 2014, 316 pages. Revue internationale de l'économie sociale: recma, (334), 122-124.
4. Bloomberg. "Big Money Backs Tiny Loans That Lead to Debt, Despair and Even Suicide." Bloomberg.com, May 3, 2022. https://www.bloomberg.com/graphics/2022-microfinance-banks-profit-off-developing-world/.
5. Clemens, Michael A., and Gabriel Demombynes. "When Does Rigorous Impact Evaluation Make a Difference? The Case of the Millennium Villages." Journal of Development Effectiveness 3, no. 3 (2011): 305–339. https://doi.org/10.1080/19439342.2011.587017.
6. Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín. "Technology and Child Development: Evidence from the One Laptop per Child Program." American Economic Journal: Applied Economics 9, no. 3 (2017): 295–320. https://doi.org/10.1257/app.20150385.
7. Evaluation Hub. "Run Evaluations." https://www.bitevaluationhub.com/run-evaluations.
8. John, B. (2024, November 14). Challenges and limitations of microfinance in achieving large-scale poverty reduction and job creation [Working paper].
9. J-PAL. "Design and Iterate the Implementation Strategy." https://www.povertyactionlab.org/resource/design-and-iterate-implementation-strategy.
10. J-PAL. "Ethical Conduct of Randomized Evaluations." https://www.povertyactionlab.org/resource/ethical-conduct-randomized-evaluations.
11. J-PAL. "Impact Evaluation Methods Table." https://www.povertyactionlab.org/sites/default/files/research-resources/impact-evaluation-methods-table.pdf.
12. J-PAL. "Power Calculations Exercise." https://www.povertyactionlab.org/sites/default/files/Exercise-PowerCalcs_0.pdf.
13. J-PAL. "Questionnaire Piloting." https://www.povertyactionlab.org/resource/questionnaire-piloting.
14. J-PAL. "Data Security Procedures for Researchers." https://www.povertyactionlab.org/resource/data-security-procedures-researchers.
15. J-PAL and IPA. Implementing Impact Evaluations: Case Study. 2023. https://poverty-action.org/sites/default/files/2023-03/Case-Study-Implementing-Impact-Evaluations.pdf.

16. Karlan, Dean, and Jacob Appel. More Than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy. Penguin, 2011.

17. NYU Office of Research. "IRB Decision Tree." https://www.nyu.edu/content/dam/nyu/research/documents/IRB/IRBDecisionTree.pdf.

18. Shelly, Sarah, et al. "Improving Communication with Participants in Behavioural Trials." 2023.

19. UNICEF. An Evaluation of the PlayPump® Water System as an Appropriate Technology for Water, Sanitation and Hygiene Programmes. 2007. http://www-tc.pbs.org/frontlineworld/stories/southernafrica904/flash/pdf/unicef_pp_report.pdf.

20. UNICEF. Ethical Considerations When Applying Behavioural Science to Programmes with Children. Innocenti, 2021. https://www.unicef.org/innocenti/media/5186/file/UNICEF-Ethical-Considerations-Behavioural-Science-Children-2021.pdf.

21. UNICEF. "UNICEF Procedure for Ethical Standards in Research, Evaluation, Data Collection and Analysis." https://www.unicef.org/evaluation/documents/unicef-procedure-ethical-standards-research-evaluation-data-collection-and-analysis.

22. UK What Works Evaluation Hub. "Pilot Impact Studies." https://evaluationhub.eif.org.uk/pilot-impact-studies/.

23. University of California Santa Barbara Library. "Data Evaluation Checklist." https://www.library.ucsb.edu/sites/default/files/attachments/data-curation/resources/DataEvaluationChecklist.pdf.

24. World Health Organization. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies. 2010. https://iris.who.int/bitstream/handle/10665/44708/9789241502320_eng.pdf.

25. Australian Institute of Family Studies. Process Evaluation. 2025. https://aifs.gov.au/sites/default/files/2025-03/2502%20EES%20process%20evaluation_1.pdf.

26. Tableau. "What Is Data Cleaning?" https://www.tableau.com/learn/articles/what-is-data-cleaning.

27. BetterEvaluation. "Data Cleaning." https://www.betterevaluation.org/methods-approaches/methods/data-cleaning.

Scale

Welcome to *Scale*!

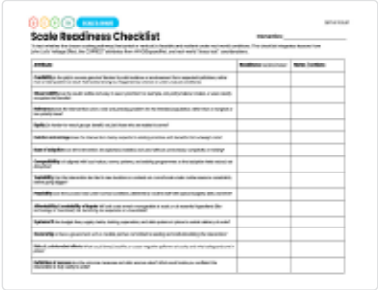
This phase explores how to expand a tested intervention, either by reaching larger populations, adapting to new regions, or becoming a part of policies and systems.

The goal of scaling an intervention is simple: **expand its reach and adapt to new contexts without losing effectiveness.** This is accomplished by identifying the essential components that drive an intervention's effectiveness, along with implementing supportive infrastructure and securing the resources necessary for sustained delivery.

Key activities in this phase include:

Scale Decision Tree

Deciding whether to expand the intervention horizontally (to new geographies or groups) and/or vertically (by integrating it into policies and systems) using a **Scale Decision Tree**



Scale Readiness Checklist

Conducting a **Scale Readiness Checklist** to assess the intervention's readiness to scale

Scale Plan

Preparing a **Scale Plan** that outlines the steps, roles, responsibilities, and financials of scale

Steps in the *Scale* phase

1 → 2 → 3

Decide Pathway

Choose the best pathway for scaling the intervention and evaluate readiness to use that pathway to deliver the intervention at scale.

TOOLS:

Scale Decision Tree

Scale Readiness Checklist

Build Support

With the problem and outcomes defined, identify the constituent behaviours. Select one that has the highest potential impact and feasibility of being changed

TOOLS:

Stakeholder Support Matrix

Plan for Scale

Prepare the people and institutions who will drive the process. identify the most influential actors, define their roles, and mapping out when and how to engage them.

TOOLS:

Horizontal Scaling Plan

Vertical Scaling Plan

Why Scale?

Without careful planning, even the most promising interventions can stall, collapse, or cause harm when spread too widely or too quickly. This section highlights why scaling matters, the risks of scaling without adaptation, the challenge of sustainability and systems integration, and the potential when scaling is done well.

The importance of scaling

For UNICEF, scaling entails the deliberate effort to ensure proven solutions benefit children and communities at the population level, not just in pilot sites. As pilots often operate under special conditions that are not necessarily present at scale, scaling is not simple replication but a guided and strategic process that requires systematic planning, and the careful consideration of institutional capacity and sustainability.

Sharing is part of this process. Through the intentional dissemination of insights, evidence, and lessons, other communities, partners, or policymakers can adopt, adapt, and multiply the impact in their own contexts.[1] This process ensures that insights can influence real world practice and policy, even where the full replication of an intervention is not possible.

However, evidence of effectiveness during pilot testing does not guarantee success at scale. Consider the following factors:

- **Feasibility**: whether the intervention is affordable and possible to be integrated with routine processes
- **Strategy**: whether the intervention delivers clear value to key stakeholders, such as the government
- **Adaptability**: whether the intervention is flexible enough to fit into policies and organizational contexts
- **Acceptability**: whether the intervention is trusted and welcomed by the target population

Even when these criteria are met, many other factors are critical in whether an intervention succeeds at scale, including the ownership and expertise of implementing teams, delivery capacity, political commitment, financing models, and alignment with policy windows.

Ideally, scaling should be considered from the onset of the DEPTHS process, to ensure that interventions are designed for real-world systems and future adaptation. However, it's also essential to resist the pressure to scale too soon, before an intervention's feasibility and effectiveness have been established. Scaling depends as much on systems, relationships, and timing as it does on evidence. Recognizing both intervention and non-intervention factors that impact scalability from the outset increases the likelihood that innovations move beyond the pilot stage to become sustainable, institutionalized solutions.[2]

1 World Health Organization (2010). Nine steps for developing a scaling-up strategy. Geneva: World Health Organization. Available at: https://iris.who.int/bitstream/handle/10665/44432/9789241500319_eng.pdf?sequence=1 [Accessed 26 Aug. 2025].

2 World Health Organization (2010). Nine steps for developing a scaling-up strategy. Geneva: World Health Organization. Available at: https://iris.who.int/bitstream/handle/10665/44432/9789241500319_eng.pdf?sequence=1 [Accessed 26 Aug. 2025].

1. The risks of scaling without adaptation and support

Not every idea that succeeds in a pilot setting will produce impact elsewhere if differences in context, systems, or delivery capacity are ignored. Scaling requires both adaptation to new environments (**horizontal scale**) and support within existing systems (**vertical scale**).

Horizontal scale — ignoring contextual differences

Community-Led Total Sanitation (CLTS) is an approach that mobilizes entire communities to end open defecation through collective action, peer pressure, and local pride — rather than through subsidies or external enforcement. The model proved highly effective in Bangladesh, where strong community cohesion enabled rapid behaviour change and widespread adoption of household latrines.

Inspired by this success, governments and partners scaled CLTS across Africa and Asia. Yet results varied when introduced in settings with weaker social cohesion, limited access to materials, or less intensive facilitation. For example, large-scale evaluations in Mali[3] and Indonesia[4] found that sanitation gains were modest or diminished over time when local enabling conditions — such as collective motivation and follow-up support — were lacking. This example shows that when community-based models are scaled without adjusting to the social dynamics and resource conditions of new settings, their initial success can quickly fade.

Vertical scale — neglecting system support

One example comes from treating malaria: Artemisinin-based therapies, adopted in the early 2000s, revolutionized care, yet rapid scale-up without strong regulatory systems allowed substandard and counterfeit drugs to circulate widely, fueling drug resistance to Artemisinin and threatening global progress.[5],[6] This example underscores that scaling biomedical innovations without investing in quality assurance, regulation, and institutional accountability can undermine both health outcomes and public trust.

Social and behaviour change programmes often face similar challenges. For example, community health workers (CHWs) played a pivotal role in raising demand for childhood vaccination through face-to-face outreach in Nepal, Senegal, and Zambia. Yet when many of these programmes expanded without sustained investments in ongoing CHW training, supervision, and/or specific adaptations tailored to the country, the impact of the programmes decreased — demonstrating that sustainable scale requires resilient, context-sensitive support systems, not just replication.[7]

Scaling without embedding interventions into systems of financing, maintenance, and accountability may produce short term impact, but it doesn't ensure sustainability in the long run. Success at scale is never about replication alone; this process requires an adaptation to context, investment in delivery systems, and integration into structures that can sustain results over time.

3 Pickering, A. J., Djebbari, H., Lopez, C., Coulibaly, M., & Alzua, M. L. (2015). Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: A cluster-randomised controlled trial. The Lancet Global Health, 3(11), e701–e711. https://doi.org/10.1016/S2214-109X(15)00144-8

4 Cameron, L., Olivia, S., & Shah, M. (2019). Scaling up sanitation: Evidence from an RCT in Indonesia. Journal of Development Economics, 138, 1–16. https://doi.org/10.1016/j.jdeveco.2018.12.001

5 Karunamoorthi K. The counterfeit anti-malarial is a crime against humanity: a systematic review of the scientific evidence. Malar J. 2014 Jun 2;13:209. doi: 10.1186/1475-2875-13-209. PMID: 24888370; PMCID: PMC4064812.

6 Newton PN, McGready R, Fernandez F, Green MD, Sunjio M, Bruneton C, et al. (2006) Manslaughter by Fake Artesunate in Asia—Will Africa Be Next? PLoS Med 3(6): e197. https://doi.org/10.1371/journal.pmed.0030197

7 Ogutu, E. A., Ellis, A. S., Hester, K. A., Rodriguez, K., Sakas, Z., Jaishwal, C., Yang, C., Dixit, S., Bose, A. S., Sarr, M., Kilembe, W., Bednarczyk, R., & Freeman, M. C. (2024). Success in vaccination programming through community health workers: a qualitative analysis of interviews and focus group discussions from Nepal, Senegal and Zambia. BMJ open, 14(4), e079358. https://doi.org/10.1136/bmjopen-2023-079358

2. Why sharing is an integral part of scaling

Sharing allows others to adopt and adapt solutions in ways that take insights further, even when full-scale replication is challenging or impossible.

Water chlorination, recognized as one of the most effective ways to make drinking water safe, is another example. When piloted in the early 2000s through the distribution of small chlorine packets for household use, early results were encouraging: diarrhoeal disease declined, households reported fewer child illnesses, and many valued the convenience of a home-based solution.

But when these programmes attempted a rapid scale-up, uptake stalled. A systematic review of point-of-use chlorination interventions found that adoption rates typically hovered around 47%, with steep declines over time, and only improved when household engagement with health workers was frequent.[8] The taste of chlorinated water was off-putting, habits around water storage were deeply entrenched, and consistent use required behaviour changes that were not easily supported in daily life. Barriers in affordability and weak distribution systems further undermined its sustained use.

However, the evidence from these pilots proved invaluable. NGOs and governments shared their findings widely, not only on the product itself, but on the behavioural barriers households faced. These insights shifted strategies beyond the sachets toward a broader mix of solutions to meet community needs and behavioural patterns: bulk chlorination at community water points, integration of safe water messaging into health worker visits, and system-level treatment at municipal plants.[9] By sharing behavioural insights alongside technical lessons, the pilots sparked innovations that ultimately reached a greater number of households.

3. When scale succeeds

Scaling done well can transform public health, with one clear example being the global effort to eradicate polio. In the mid-20th century, polio was one of the most feared diseases worldwide, paralysing hundreds of thousands of children each year. Early vaccines proved highly effective in small-scale trials, but success at the population level required more than the product alone: it demanded coordinated systems, global partnerships, and community trust.

Scaling polio vaccination was a dual effort. On the systems side, the WHO, UNICEF, governments, and partners built vast delivery platforms, from cold-chain logistics and mass immunization campaigns to surveillance networks that could quickly detect and respond to outbreaks. On the behavioural side, trusted health workers and community leaders carried out house-to-house outreach, addressing fears, misinformation, and resistance. In many settings, sustained advocacy from religious leaders and local champions was crucial to achieving acceptance.

The results are historic. Polio cases have fallen by more than 99% since 1988, and the disease is now close to eradication. This success was not automatic: it came from pairing evidence-based tools with deliberate system integration, political commitment, and behavioural strategies that supported uptake.

8 Crider, Yoshika & Tsuchiya, Miki & Mukundwa, Magnifique & Ray, Isha & Pickering, Amy. (2023). Adoption of Point-of-Use Chlorination for Household Drinking Water Treatment: A Systematic Review. Environmental Health Perspectives. 131. 10.1289/EHP10839.-

9 Crider, Yoshika & Tsuchiya, Miki & Mukundwa, Magnifique & Ray, Isha & Pickering, Amy. (2023). Adoption of Point-of-Use Chlorination for Household Drinking Water Treatment: A Systematic Review. Environmental Health Perspectives. 131. 10.1289/EHP10839.

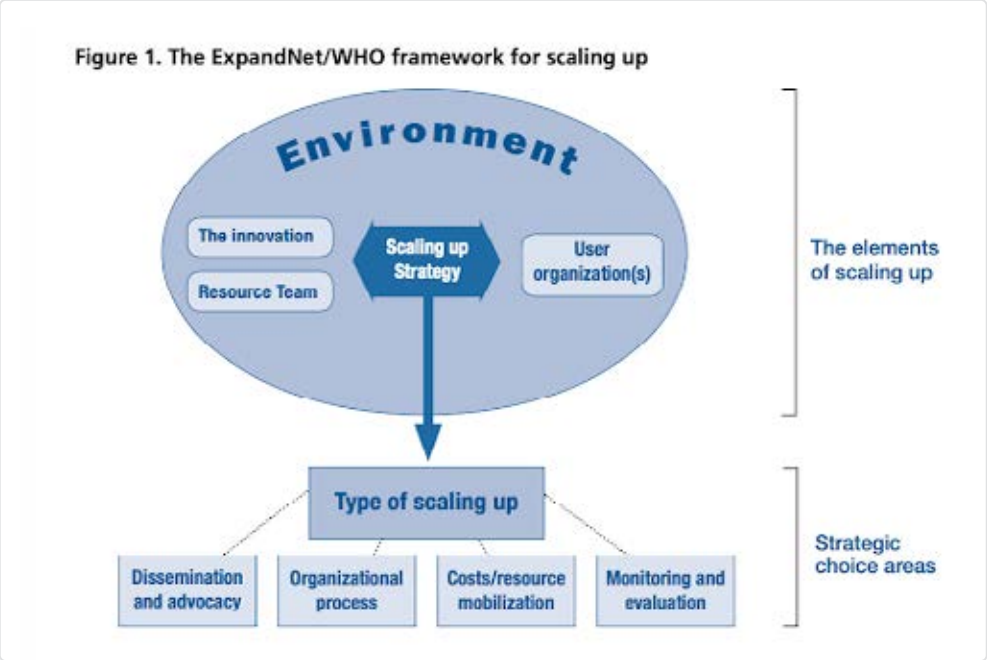
How can we Scale?

Just as designing and testing interventions requires structured methods, scaling their impact calls for dedicated strategies and tools. This chapter is based on the Nine Steps for Developing a Scaling-Up Strategy by ExpandNet/WHO (2010). The framework has been adapted to support behavioural interventions.

Common pitfalls

Throughout the Scale phase, keep these possible issues in mind:

- **Weak stakeholder engagement.** Scaling needs strong support from governments, partners, and staff delivering the intervention. Building coalitions, aligning incentives, and negotiating roles up front can mitigate the risk of failure.
- **Underestimating costs and resources needed.** Scaling often costs more than expected. Striving to serve communities adds time and effort, and beyond the pilot phase, both human and financial resources are often not budgeted for. If costs are underestimated, programmes cannot be sustained.
- **Growing complexity and short term funding.** As interventions scale, extra tools, training or technology are often added, particularly if new partners or funders add tools based on their own priorities — making programmes more difficult for both workers to deliver and systems to manage. Many scale efforts rely on project money or temporary donor support; as such, when this funding ends, programmes often shrink or come to a halt, even if they are succeeding. This unfortunately occurs when long term financing and government budgets are not built in from the start.



ExpandNet. n.d. "Scaling-Up Framework and Principles." Accessed February 17, 2026. https://expandnet.net/scaling-up-framework-and-principles/

The case of increasing childhood vaccination uptake in Lebanon

In the previous chapter, *Test Hypotheses*, the project team ran a randomized controlled trial (RCT) of the reminder postcard and found it increased vaccination uptake by 6.7 percentage points compared to the control group. While the evidence of effectiveness was clear, the team still needed to confirm that the intervention was truly ready to scale.

The Ministry of Public Health (MoPH), a key partner, was ready to secure the intervention, and there was strong demand from other districts not included in the pilot. These factors pointed to a dual approach to scale: horizontal expansion (scaling the intervention to new regions) alongside vertical integration (embedding the postcard into the MoPH's national health systems). While the evaluation results were encouraging, the team was aware that scaling would require more than evidence of effectiveness.

To mitigate any risks during scale, the team worked through a checklist of good practices and signals to review before scaling. This helped them to identify several key considerations. For example, the RCT results and subgroup analysis gave the intervention credibility, demonstrating that the approach could improve equity by benefiting low income and refugee families at a lower cost than existing practices. However, they also identified a key risk: systems fit was only partial. While easy to use, the postcards were not yet integrated into routine procurement or supply chains, creating a risk of stockouts at scale.

To address this risk, the team returned to the Stakeholder Map created in the Define phase, adding new stakeholders that would be key to scaling success. With this updated list, they planned the engagement approach, mapped interests, and identified what each actor needed to know and do. For example, the MoPH required cost benefit data to approve a budget line, health centre supervisors needed simple training materials, and donors needed evidence of a sustainable exit strategy.

The team also ensured donors would fund the government during the transition, until systems were in place to integrate the postcard into routine supply chains within vaccination policy.

Finally, the team prepared plans for both geographical and population expansion (horizontal scale) and integration with the MoPH's system (vertical scale). The Horizontal Scaling Plan detailed a phased rollout to three new regions, leveraging existing vaccine supply chains for distribution and using a simple dashboard to monitor fidelity, equity, and cost effectiveness. In parallel, the Vertical Scaling Plan outlined steps for institutionalization: the MoPH issuing a circular (an official document formalizing the intervention's integration into health services), a permanent budget allocated for printing, and inclusion of the postcard in official supply order forms and training modules.

By systematically deciding on a pathway, checking readiness, building support, and creating detailed plans, the team could strategically scale the intervention, expanding its impact while embedding it into the systems that would sustain it for years to come.

***Note:** While this is a real project that closely followed a very similar process to DEPTHS, there were a few tools from the toolkit that the project team did not apply during implementation. In those cases, we've gone back and retrospectively applied the tools using real project data to illustrate how they might have looked if they had been used at the time.*

STEP 1:
Decide pathway

Associated Tools:

- Scale Decision Tree
- Scale Readiness Checklist

In this step:

Here, the focus is on two key tasks:

1. **Choosing the pathway for scaling.** Use the Scale Decision Tree to reflect on the most appropriate scaling approach for the intervention, including the following:
    - Horizontal expansion, where interventions are scaled to new geographies and/or groups.
    - Vertical integration, where the intervention is embedded into existing policies, budgets, and systems.

    If there is a lack of political and community demand, evidence is weak, or results consistently fail to show value or feasibility, stop the scaling process and strengthen the intervention.

2. **Checking readiness.** Once the decision on what type of scaling approach has been made, use the Scale Readiness Checklist to reflect on whether the intervention is likely to remain effective when delivered at larger scale and under routine system conditions.

Why it matters:

Jumping into scale without first reflecting on the most suitable approach can push teams down the wrong path and waste time and resources.

In addition to positive evaluation results, other strong predictors of success include financing, delivery capacity, and political support. For example, a mobile vaccination team may succeed in one city but struggle in remote areas, if fuel costs are high or staff cannot travel regularly.

How to do it:

1. Reflect on the best pathway for scaling

a. The first step is to decide whether scaling the intervention is worthwhile — if so, whether a horizontal or vertical pathway is most appropriate. Both may be desirable over time, but identifying a primary focus early on helps to concentrate effort for maximum impact.

- **Horizontal scaling** involves extending an intervention to new places or groups, while ensuring that it is suited to local conditions and continues to preserve the core behavioural elements that drive its impact. This can also be referred to as expansion or replication.
- **Vertical scaling** refers to embedding an intervention into higher-level systems — such as policies, regulations, and institutional frameworks — to ensure it becomes formally adopted and supported at scale. For example, an intervention that improves service delivery may be written into national guidelines or included in government budgets for it to continue without external support. Vertical scaling typically depends on government leadership and long-term commitment, as well as stable funding and policy alignment.

Most teams already have an idea of which path they want to pursue. The Scale Decision Tree helps to test this tentative direction for scale by posing clear questions about evidence, government backing, policy opportunities, and user demand.

Begin with reviewing findings from the Test Hypotheses phase. Scaling should be considered only when evaluation results demonstrate that the intervention has a significant positive impact on the outcomes outlined in the Theory of Change from the Prototype Designs phase. Three criteria are needed to assess whether the evidence supports scaling an intervention:

- **Effectiveness:** Evidence from the Test Hypotheses phase shows measurable improvements in the behaviour targeted by the intervention. This may come from different types of evidence: in some cases, rigorous trials such as RCTs or quasi-experimental designs provide statistically significant results; in others, strong observational evidence shows consistent patterns across sites, or large-scale monitoring data reveals improvements in coverage and demand.
- Effectiveness can also be established when different data sources converge to show meaningful impact, or when results are considered 'good enough' to justify adoption at scale given cost, feasibility, and demand. For example, a text message reminder for contraceptive use was tested across multiple clinics and consistently improved adherence, with results holding across different settings, not only one site.

- **Equity:** Building on effectiveness results, consider whether the intervention benefits vulnerable populations, not only those most easily reached. Sub-group analysis of pilot results can help to assess this – for example, by disaggregating outcomes across gender, age, geography, or socioeconomic status. In some cases, such analysis may not be feasible, and judgments about equity will rely on more subjective or contextual assessments — such as whether the intervention addresses known barriers faced by marginalized groups.

  For example, contraceptive reminder messages may be assessed not only for their impact among urban women with stable phone access, but also for their relevance and accessibility to adolescents and rural populations who face greater barriers.

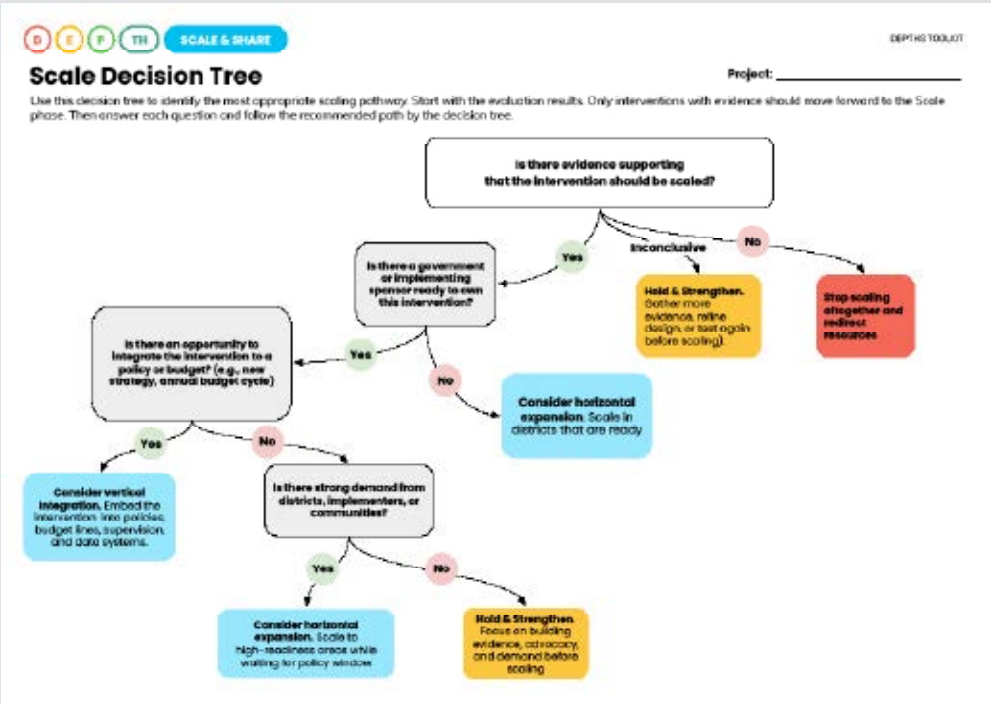
- **Affordability:** Cost–benefit analysis from the pilot should demonstrate that the intervention's costs are proportionate to its impact, and manageable within routine budgets. For example, bulk SMS reminders may cost only a few cents per user per month, making it realistic to integrate them into existing national communication platforms.

b. If the evaluation reveals no improvement in behavioural outcomes, benefits limited to the easiest-to-reach groups, or costs that exceed what systems can realistically sustain, scaling the intervention is not justified. In such cases, stop the scaling process and focus on assessing the lessons learned.

c. If the intervention does not meet all three criteria, the evidence is not strong enough to justify scaling. In these situations, scaling should be paused. Consider refining the intervention, gathering additional data, or conducting another round of testing before moving forward.

d. When pilot data indicates that an intervention is effective, equitable, and affordable, scaling is justified. The choice of scaling pathway, however, depends on the context and prevailing conditions. To decide which pathway to follow, consider the following three questions:

- **Is there a government or institutional sponsor ready to take ownership?** If yes, the next step is to look for opportunities to integrate the intervention into policy, budgets, or delivery systems.



This is the pathway to vertical scaling. For example, a contraceptive reminder service could be built into the national digital health platform or funded through government family planning budgets.

- **If no sponsor is ready, is there strong demand from districts, implementers, or communities?** If local demand is strong, horizontal expansion may be the best first step. The intervention can expand reach in high-readiness areas while continuing to advocate for government support. For example, clinics in certain districts may already have the staff and resources to roll out contraceptive reminder messages quickly.
- **If neither ownership nor demand is clear, what can be strengthened now?** In this case, it is better to pause and focus on building the enabling conditions. This may mean advocacy to create political will, further testing to adapt the design, or working with partners to build capacity and interest. For example, if contraceptive reminder messages have not yet attracted government sponsorship or community demand, consider pausing scaling while gathering more evidence, tailoring the design, and engaging stakeholders to build support.

The most appropriate scaling pathway will be identified by the end of the Scale Decision Tree, leading to the next step: reviewing the intervention's readiness for pursuing the chosen pathway.

2. Assess readiness

Once a pathway for scale is established, the next step is to rigorously assess the intervention's preparedness to move in that direction. Assessing readiness requires systematically testing the intervention against key conditions, identifying which are firmly in place, where critical gaps remain, and what investments or adaptations will be necessary to enable scale.

The Scale Readiness Checklist helps with this process. It draws on three complementary frameworks:

- **John List's** Voltage Effect framework[10]: This framework explores why promising pilots often fail at scale, such as dependence on special conditions or underestimating real-world costs.
- **CORRECT attributes** (WHO/ExpandNet[11]): This framework establishes practical criteria for scalability, including credibility, relevance, ease of adoption, and compatibility with existing systems.
- **Option C thinking**[12]: This framework stress-tests whether success can be maintained when budgets are tighter, systems are overstretched, or delivery conditions differ from the pilot.

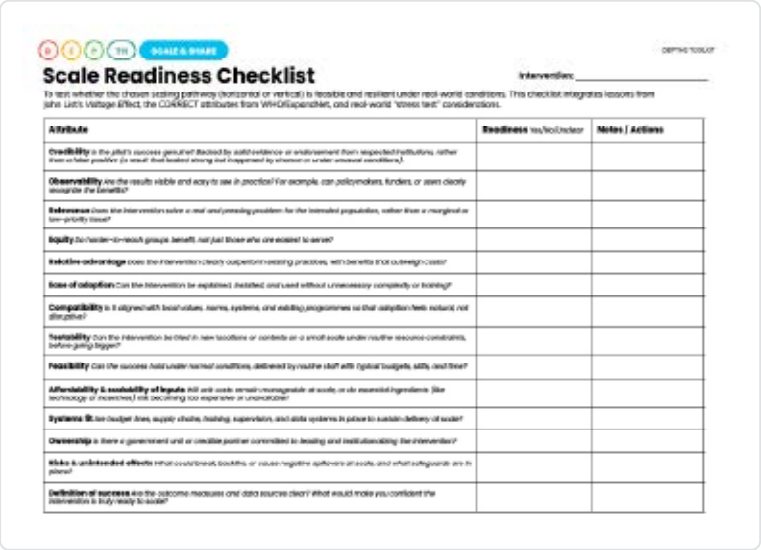
Before completing the checklist, review the operational dimensions that determine whether an intervention can succeed under routine conditions. Pilots often benefit from extra resources or ideal circumstances that cannot be assumed at scale. Appendix 2 outlines the areas most likely to falter during expansion, such as staffing and supervision, supply chains, financing, and community access. Considering these dimensions first will provide a more realistic foundation for completing the Scale Readiness Checklist, helping to spot hidden risks and identify elements that need adaptation or reinforcement.

To complete the Scale Readiness Checklist:

a. Start by reviewing Table 2 below, which details why each readiness attribute matters, the risks if it is not met, and possible actions to address gaps.

b. For each attribute listed in the Scale Readiness Checklist, decide whether the intervention currently meets the requirement for scale. Record your judgment as Yes (ready), No (not ready), or Unclear (evidence is mixed or missing).

c. In the Notes / Actions column, describe what is missing and what steps could strengthen the intervention before scaling.



Scale Readiness Checklist

Intervention: _____

Attribute	Readiness Yes/No/Unclear	Notes / Actions
Credibility Is the pilot's success genuine? Backed by solid evidence or endorsement from respected institutions, rather than a false positive (a result that looked strong but happened by chance or under unusual conditions).		
Observability Are the results visible and easy to see in practice? For example, can policymakers, funders, or users clearly recognize the benefits?		
Relevance Does the intervention solve a real and pressing problem for the intended population, rather than a marginal or low-priority issue?		
Equity Do harder-to-reach groups benefit, not just those who are easiest to serve?		
Relative advantage Does the intervention clearly outperform existing practices, with benefits that outweigh costs?		
Ease of adoption Can the intervention be explained, installed and used without unnecessary complexity or training?		
Compatibility Is it aligned with local values, norms, systems, and existing programmes so that adoption feels natural, not disruptive?		
Testability Can the intervention be tried in new locations or contexts on a small scale under routine resource constraints, before going bigger?		
Feasibility Can the success hold under typical conditions, delivered by routine staff with typical budgets, skills, and time?		
Affordability & availability of inputs Will unit costs remain manageable at scale, or do essential ingredients (like technology or incentives) risk becoming too expensive or unavailable?		
Systems fit Are budget lines, supply chains, training, supervision, and data systems in place to sustain delivery at scale?		
Ownership Is there a government and or credible partner committed to leading and institutionalizing the intervention?		
Risks & unintended effects What could break, backfire, or cause negative spillovers at scale, and what safeguards are in place?		
Definition of success Are the outcome measures and data sources clear? What would make you confident the intervention is truly ready to scale?		

10 For more information on John List's Voltage Effect framework, see Appendix 1 at the end of the chapter.
11 World Health Organization & ExpandNet. Nine Steps for Developing a Scaling-Up Strategy. WHO, 2010.
12 Al-Ubaydli, Omar, and John A. List. "Will It Scale?" Issues in Science and Technology 41, no. 1 (Fall 2024): 34–36.

TABLE 2. ATTRIBUTES FOR THE SCALE READINESS CHECKLIST

ATTRIBUTE	DESCRIPTION	EXAMPLE	POTENTIAL MITIGATION ACTIONS
Credibility: Is the intervention's success genuine, backed by reliable evidence or respected endorsements?	John List's *Voltage Effect* warns that many pilots risk being "false positives" — appearing successful under special conditions but failing when scaled. As discussed in the Test Hypotheses chapter, a false positive occurs when an evaluation suggests an effect that is not truly there, much like a pregnancy test showing a positive result when the person is not pregnant.	The Climate Schools programme was an online school-based intervention using interactive lessons and cartoon scenarios to prevent depression and anxiety in adolescents. Small early studies showed promising results. But when tested in a large trial across 18 schools, it did not improve core mental health outcomes, and in some measures students in intervention schools did worse than controls. This highlights the credibility risk: early positive results may not hold up under routine conditions or larger, more diverse settings.[13]	Confirm pilot results are statistically robust and not dependent on exceptional circumstances (e.g., "impact observed only in one district with unusually high staff support"). If these checks were not already completed during the Test Hypotheses phase, they should include multiple-comparison adjustments, a review of effect sizes and confidence intervals, and analyses of sensitivity or robustness.
Observability: Can the benefits of the intervention be seen and recognized by others?	Observability, one of the CORRECT attributes identified by WHO/ExpandNet, refers to whether the benefits of an intervention can be seen and recognized by others. Interventions are more likely to scale when positive outcomes are visible.	In a programme with US college students, some participants saw their steps displayed on a public leaderboard, while others received only private feedback. Those in the public group walked more, showing that visible results (in this case, peers exercising) can motivate wider uptake.[14]	Assess whether the benefits of the intervention are noticeable to others. If not, make results more visible through strategies such as public demonstrations, recognition events, or presenting data in formats that communities, health workers, or leaders can easily see and understand.
Relevance: Does the intervention address a problem that feels urgent to the population?	Relevance emphasizes that scaling is most successful when the issue resonates as a clear priority for people and institutions. If the problem is not seen as important, motivation to adopt is weaker. Teams should therefore examine whether the issue is recognized as a priority and adjust framing if needed.	In Spain, the EIRA study aimed to change multiple unhealthy behaviours — smoking, physical activity, and diet — among adults aged 45–75 in primary health care centres. While the intervention improved diet, it had little impact on smoking or physical activity. One reason was that increasing physical activity was not perceived as urgent. This lack of resonance reduced motivation to engage, limiting broader uptake.[15]	Assess whether the issue is seen as a top priority by the target group and policymakers. If not, adapt how the intervention is framed. For example, connect it to existing concerns people already care about, such as protecting family health, saving money, or reducing stress.

Table continues on following page.

13 Andrews JL, Birrell L, Chapman C, Teesson M, Newton N, Allsop S, McBride N, Hides L, Andrews G, Olsen N, Mewton L, Slade T. Evaluating the effectiveness of a universal eHealth school-based prevention programme for depression and anxiety, and the moderating role of friendship network characteristics. Psychol Med. 2023 Aug;53(11):5042-5051. doi: 10.1017/S0033291722002033. Epub 2022 Jul 15. PMID: 35838377.

14 Lee JJ, Kim Y, Welk GJ, Hannon JC. The effect of using onymous and anonymous normative feedback on physical activity in college students: a randomized controlled trial. BMC Sports Sci Med Rehabil. 2020;12:27. doi:10.1186/s13102-020-00202-y.

15 Zabaleta-del-Olmo, E., Casajuana-Closas, M., López-Jiménez, T. et al. Multiple health behaviour change primary care intervention for smoking cessation, physical activity and healthy diet in adults 45 to 75 years old (EIRA study): a hybrid effectiveness-implementation cluster randomised trial. BMC Public Health 21, 2208 (2021). https://doi.org/10.1186/s12889-021-11982-4

ATTRIBUTE	DESCRIPTION	EXAMPLE	POTENTIAL MITIGATION ACTIONS
Equity: Does the intervention reach those who are underserved or at higher risk?	Scaling a solution that only works for well-resourced groups can widen gaps. Equity means ensuring that harder to reach populations also engage and benefit from an intervention, further ensuring that scaling does not deepen any inequalities.	A randomized trial of a mobile health programme for adults with type 2 diabetes in the United States showed uneven engagement. Non-white participants, those with lower health literacy, and older adults were significantly less likely to use the programme. As a result, the groups most in need were the least likely to benefit—an equity risk if the programme were scaled.[16]	Review pilot results by different groups such as age, gender, or literacy. If gaps appear, plan adaptations before scaling. This can include using simpler messages, voice calls, face-to-face support, or trusted community helpers.
Relative advantage: Does the intervention clearly outperform current practice?	People adopt innovations that save time, effort, or cost. If the benefit over current practice is weak or fades in routine use, uptake will stall. It is important to ensure that improvements are clear, lasting, and valuable to users.	In a large randomized trial in India, lab-validated improved cookstoves initially reduced smoke exposure, but the gains disappeared by year two. There were no improvements in health or fuel use because households used stoves irregularly, didn't maintain them, and usage declined over time. The stoves did not offer a compelling enough relative advantage over existing practices to sustain behaviour at scale.[17]	When reviewing relative advantage, record whether the intervention delivers clear, understandable improvements over current practice for users and institutions (e.g., fewer steps, less time, lower ongoing cost, better reliability). Note any signs that benefits fade without extra support, or that maintenance/training burdens erode day-to-day value. If the relative advantage is weak, identify adaptations that could strengthen it.
Ease of adoption: Can the intervention be used without too much training, supervision, or effort?	Interventions that are simple and convenient spread more easily. If adoption requires specialist skills, ongoing support, or tolerance for negative side effects, uptake is limited. Simplifications may be needed.	A cluster-randomized trial in Bangladesh evaluated household point-of-use chlorination for drinking water. While chlorination effectively reduced diarrhoea in the short term, sustained adoption was very low. Many households reported that dosing was confusing, the process felt burdensome, and the taste of chlorinated water discouraged consistent use. Despite health benefits, the intervention's lack of ease of adoption made it difficult to scale without major modifications.[18]	Assess whether people can use the intervention quickly and consistently under normal conditions, without ongoing external support. Note if tasks require specialist skills, repeated reminders, or the tolerance of negative side effects. If adoption barriers are high, identify simplifications or supports needed before scaling—for example, automated dosing, integration with routine services, or redesigning delivery.

Table continues on following page.

16 Nelson LA, Mulvaney SA, Gebretsadik T, Ho Y-X, Johnson KB, Osborn CY. Disparities in the use of a mHealth medication adherence promotion intervention for low-income adults with type 2 diabetes. J Am Med Inform Assoc. 2016;23(1):12-18.

17 Hanna R, Duflo E, Greenstone M. Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves. American Economic Journal: Economic Policy. 2016;8(1):80–114. doi:10.1257/pol.20140008.

18 Pickering AJ, Crider Y, Sultana S, Swarthout J, Goddard FG, Anjerul Islam S, Sen S, Ayyagari R, Luby SP. Effect of in-line drinking water chlorination at the point of collection on child diarrhoea in urban Bangladesh: a double-blind, cluster-randomised controlled trial. Lancet Glob Health. 2019 Sep;7(9):e1247-e1256. doi: 10.1016/S2214-109X(19)30315-8. PMID: 31402005.

ATTRIBUTE	DESCRIPTION	EXAMPLE	POTENTIAL MITIGATION ACTIONS
Compatibility: Is the intervention aligned with community norms and institutional routines?	Even effective interventions could fail if they clash with cultural values or established systems. Compatibility increases when the intervention fits with everyday life and existing systems.	A randomized controlled trial of household latrine promotion in rural India (the MANTRA programme) found that even when latrines were constructed, use remained low. One key reason was cultural incompatibility: open defecation was widely perceived as healthier and more convenient, and latrines were not considered acceptable spaces within household compounds. Despite financial subsidies and infrastructure, low compatibility with community norms limited adoption and undermined the intervention's impact.[19]	When reviewing compatibility, assess whether the intervention fits with community norms and institutional practices. Consider whether values, traditions, or routines might conflict with the intervention's design. If compatibility is weak, note the adaptations required before attempting to scale. For example, modifying the intervention to align with cultural norms, reframing messages to match community values, or embedding the approach into existing service delivery systems.
Testability: Can the intervention be tried in new areas under routine conditions, before full rollout?	People and institutions are more likely to adopt an intervention if they can "try before they buy." Phased rollouts and pilot sites reduce risk and allow adaptation. If large upfront investment is needed, entry points should be smaller.	In rural Kenya, chlorine dispensers were installed at communal water sources. Unlike earlier household-based approaches, dispensers allowed families to test water treatment immediately, with little cost or burden. Uptake was higher and more sustained because people could experience the benefits under routine conditions before committing to long-term use.[20]	Assess whether the unit costs observed in the pilot can be maintained or reduced at scale, and whether essential inputs (e.g. commodities, technologies, staff, and infrastructure) can be supplied consistently. If costs escalate or supply chains are fragile, explore strategies such as alternative delivery channels, pooled procurement, or integration into existing logistics systems.
Feasibility: Can routine staff deliver the intervention within normal budgets and workloads?	Pilots often benefit from extra supervision, funding, or incentives that are not sustainable. Feasibility means the intervention can succeed under routine conditions with existing staff and resources.	A review of strategies to improve provider practices across low- and middle-income countries found that many interventions relying on intensive supervision, training, or large financial incentives produced strong results in pilots, but failed to sustain impact at scale. Health systems could not replicate the support provided during trials. Feasibility depends on whether improvements can be delivered under routine budgets and staffing constraints.[21]	Note whether the intervention can be delivered by frontline staff under normal conditions (e.g., typical work load, existing training, and routine budgets). If the pilot depended on extra supervision, incentives, or donor subsidies, record these gaps. Consider what adaptations could reduce resource intensity to make the model feasible at scale. For example, simplified protocols, digital tools, or integration into existing workflows.

Table continues on following page.

19 Clasen T, Boisson S, Routray P, Torondel B, Bell M, Cumming O, Ensink J, Freeman M, Jenkins M, Odagiri M, Ray S, Sinha A, Suar M, Schmidt WP. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. Lancet Glob Health. 2014 Nov;2(11):e645-53. doi: 10.1016/S2214-109X(14)70307-9. Epub 2014 Oct 9. PMID: 25442689.

20 Kremer M, Miguel E, Mullainathan S, Null C, Zwane AP. Spring cleaning: Rural water impacts, valuation, and property rights institutions. Q J Econ. 2011;126(1):145–205. doi:10.1093/qje/qjq010.

21 Rowe AK, Rowe SY, Peters DH, Holloway KA, Chalker J, Ross-Degnan D. Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: a systematic review. Lancet Glob Health. 2018;6(11):e1163–e1175. doi:10.1016/S2214-109X(18)30398-X.

ATTRIBUTE	DESCRIPTION	EXAMPLE	POTENTIAL MITIGATION ACTIONS
Affordability and scalability of inputs: Will costs and supplies remain manageable at scale?	Inputs that are affordable in pilots may not remain so when scaled. Ensure that unit costs and supply chains are sustainable and reliable at national level.	In a multi-country evidence synthesis, a randomized controlled trial in Nepal tested the use of micronutrient powders (MNPs) to reduce childhood anaemia. While effective in small pilots, evaluations of larger rollouts revealed that procurement and distribution costs, coupled with weak supply chains, made routine delivery difficult. Stock-outs were common, and sustained financing was a challenge for government systems. The scalability of MNP sachets became the primary barrier to scale.[22]	Record whether the unit costs observed in the pilot can be maintained or reduced when scaled, and whether key inputs (commodities, technologies, staff, and infrastructure) can be supplied reliably. If costs balloon or supply chains are fragile, note these risks explicitly. Consider whether alternative delivery channels, pooled procurement, or integration into existing logistics systems could make the intervention more affordable and scalable.
Systems fit: Are the supporting systems in place to sustain delivery?	Scale requires budgets, supply chains, training, supervision, and data systems. If these are weak, even strong interventions can collapse after expansion.	In India, the large-scale Janani Suraksha Yojana (JSY) conditional cash transfer programme increased facility births but strained health system capacity. Evaluations found that many facilities lacked sufficient staff, drugs, and supplies to handle the surge. While uptake was high, the absence of parallel investments in health system infrastructure and supervision undermined quality of care and limited the programme's overall impact.[23]	Record whether key systems (budgets, logistics, workforce training, supervision, and data monitoring) are strong enough to support delivery at scale. If gaps exist, note them clearly and identify what investments or partnerships are needed. Without alignment with supporting systems, scaling risks short-term expansion followed by breakdown.
Ownership: Is there strong commitment from governments, institutions, and communities?	Without ownership, interventions often collapse once external support ends. Ownership means leaders, staff, and communities view the intervention as theirs, and commit resources to sustaining it.	In Zambia, a randomized evaluation of a performance-based financing (PBF) scheme showed that while the pilot improved service delivery indicators, government ownership was limited. The programme was largely donor-driven, and when external funding and technical support ended, many gains were not sustained.[24]	Note whether governments and communities see the intervention as theirs, which can be reflected in political commitment, the allocation of domestic resources, and integration into existing institutions. If ownership is weak, note whether advocacy, alignment with national priorities, or participatory co-design processes are needed before scaling.

Table continues on following page.

22 Locks LM, Reerink I, Hedlund K, Peña-Rosas JP, Jefferds ME, Mclean MS, et al. Micronutrient powder programs: lessons learned for integrated infant and young child feeding. Am J Clin Nutr. 2017;105(5):1126–1136. doi:10.3945/ajcn.116.144055.

23 Lim SS, Dandona L, Hoisington JA, James SL, Hogan MC, Gakidou E. India's Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities: an impact evaluation. Lancet. 2010;375(9730):2009–2023. doi:10.1016/S0140-6736(10)60744-1.

24 Friedman J, Qamruddin J, Chansa C, Das AK, McMahon S, McPake B. Impact evaluation of Zambia's health results-based financing pilot project. Health Policy Plan. 2016;31(9):1117–1124. doi:10.1093/heapol/czw049.

ATTRIBUTE	DESCRIPTION	EXAMPLE	POTENTIAL MITIGATION ACTIONS
Risks and unintended effects: What could go wrong once scaled?	Scale can create problems such as system overload, inequities, or negative spillovers. Anticipating risks allows for the design of safeguards.	Analysis of seven scaled physical activity and nutrition interventions in Australia found that even when programmes achieved their intended outcomes, scale sometimes produced unintended consequences. These included increased workload for frontline staff, reduced sustainability in some settings, and shifting priorities that undermined programme value. The study concluded that mechanisms of scale can generate both positive and negative effects, and that anticipating risks is vital.[25]	Explore what could go wrong if the intervention is scaled—for example, service overload, inequities, or unintended spillovers. Examine evidence from the pilot, or from similar interventions where such risks have occurred, and record whether safeguards are in place, such as phased rollouts, monitoring systems, or equity tracking. If risks are likely, scaling plans should incorporate mitigation strategies, rather than assume that expansion will simply multiply benefits.
Definition of success: Are outcomes and data sources clear and measurable?	Without a clear definition of success, it's difficult to know if scale is working. Outcomes should be specific, credible, and meaningful to communities, governments, and funders.	In Kenya, the WelTel Kenya1 randomized controlled trial used weekly SMS check-ins to support antiretroviral therapy adherence. Unlike many mHealth pilots that track only engagement (e.g., number of messages sent), this trial defined success in terms of biologically verified adherence and viral suppression, measured through electronic monitoring and clinical outcomes. By using clear, rigorous outcome definitions, the trial gave funders and policymakers confidence that the intervention's effects were real and worth considering for broader investment.[26]	Record whether outcomes of interest are explicitly defined and tied to credible data sources. Ask: What would success look like at scale, and how will it be measured? If outcomes are vague (e.g., "improved awareness"), note the need to sharpen definitions. Where possible, ensure indicators are linked to behaviour change or health outcomes, not just process measures, so teams and funders can track whether scaling is truly delivering value.

25 Koorts H, Eakin E, Estabrooks P, Timperio A, Salmon J, Bauman A. Mechanisms of scaling up: combining a realist perspective and systems analysis to understand successfully scaled interventions. Int J Behav Nutr Phys Act. 2021;18:61. doi:10.1186/s12966-021-01103-0.

26 Lester RT, et al. Lancet. 2010;376(9755):1838-1845. doi:10.1016/S0140-6736(10)61997-6.

Case study: increasing childhood vaccination uptake in Lebanon

The tools mentioned in this step of the Scale phase were not used by the original project team. This case study is a recreated example based on real project data and context.

The decision

After reviewing the evaluation results, the team concluded that both horizontal and vertical scaling were possible, however, the immediate priority was horizontal expansion. Several districts outside of the pilot area had already expressed interest, and the postcard intervention was simple and low-cost to replicate. At the same time, the team decided to prepare for vertical integration into the Ministry of Public Health's Expanded Programme on immunization (EPI), which would require policy and budget adjustments. This dual approach balanced short-term expansion with longer-term sustainability.

Checking readiness

With the decision clear, the team applied the Scale Readiness Checklist to test whether the postcards could hold up under scale.

D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Scale Readiness Checklist

Intervention: Appointment reminder card

Attribute	Readiness Yes/No/Unclear	Notes / Actions
Credibility	Yes	The RCT showed a significant 6.7 percentage point increase in timely vaccination. Results were statistically robust. Action: Package findings into briefs and presentations endorsed by trusted institutions (e.g. Ministry of Public Health, UNICEF) to strengthen credibility.
Observability	Yes	Caregivers bring postcards to clinics, which staff can see. Benefits are tangible and easy to demonstrate. Action: Encourage staff to share stories and examples with district managers to reinforce visibility.
Relevance	Yes	Childhood vaccination is a national health priority, especially given pockets of under-immunisation among vulnerable groups. Action: Frame postcards as a low-cost way to meet immunisation targets.
Equity	Yes	Analysis showed positive effects even among low-income families and refugee households. Action: Monitor distribution to ensure marginalised groups continue to receive cards equitably.
Relative advantage	Yes	Postcards provide a reliable reminder compared to informal word of mouth. They are inexpensive, physical, and hard to miss. Action: Collect testimonials from caregivers to highlight added value over existing practices.
Ease of adoption	Mostly yes	Health workers reported postcards were simple to distribute and fill in, though training was needed for consistency. Action: Create a pictorial job aid for health workers to reduce variation.
Compatibility	Yes	Postcards fit naturally into clinic workflows and community norms. Caregivers are familiar with keeping health documents at home. Action: Link postcards to existing vaccination cards for smoother integration.
Testability	Yes	Districts can adopt postcards with minimal investment and test under routine conditions. Action: Plan phased introduction in 3–4 districts before national rollout.
Feasibility	Mostly yes	Clinics can manage postcard distribution under current staffing, but supervision may be inconsistent. Action: Add a tick-box in supervisory checklists to monitor use.
Affordability & scalability of inputs	Yes	Printing costs are low (less than $0.20 per postcard). Cards can be bundled with vaccine shipments. Action: Secure framework agreements with local printers to keep costs stable.
Systems fit	Partly	Logistics and supply systems can handle postcard delivery, but budgets are not yet institutionalised. Action: Advocate for a dedicated budget line in the national immunisation plan.
Ownership	Unclear	Ministry of Public Health expressed interest but has not yet formalised commitment through policy or budgets. Action: Engage the EPI unit early, demonstrate results from district pilots, and prepare a circular for formal adoption.
Risks & unintended effects	Yes, some risks	Risks include stock-outs if printing is not managed well, novelty fade among caregivers, and inconsistent clinic use. Action: Plan design refreshes every 2 years and embed postcards into national supply chain codes.
Definition of success	Yes	Success is defined as increased on-time vaccination, measured through clinic records. Equity effects for vulnerable groups are tracked separately. Action: Maintain outcome measures linked to health impact, not only distribution numbers.

- **Credibility**: The randomized trial showed a 6.7 percentage point increase in on-time vaccination. The effect was consistent across sites, reducing the risk of a false positive.
- **Equity**: Sub-group analysis showed that low income and refugee families also benefited, not only higher income households. This strengthened the case for national adoption.
- **Observability and relative advantage**: The postcards were highly visible. Health workers saw caregivers bringing them to clinics, and caregivers found them easier to track than verbal reminders. This created a clear improvement over existing practice.
- **Feasibility and systems fit**: Printing and distribution were straightforward in the pilot but not yet built into national systems. Without a dedicated budget line or supply code, there was a risk of stock-outs if the programme expanded too quickly.
- **Ownership and timing**: The Ministry of Public Health expressed interest but had not formalized the intervention's adoption. The team noted the importance of aligning with the upcoming budget cycle to create an entry point for institutionalization.
- **Risks**: If the postcards were poorly managed, stock-outs could harm trust. Caregivers might also lose interest over time. The team flagged the need for design refreshes every two years and supply chain integration to mitigate these risks.
- **Definition of success**: Success was defined as not only the number of postcards distributed, but improved on-time vaccination rates, particularly among vulnerable groups.

STEP 2:
Build support

Associated Tools:

- Stakeholder Support Matrix

In this step:

Scaling depends not only on the strength of the intervention, but on the alignment, commitment, and readiness of key actors. After selecting the pathway for scale and assessing readiness, the next step is to prepare the people and institutions who will drive the process.

Begin by identifying the most influential actors, defining their roles, and mapping out when and how to engage them. Move beyond a simple stakeholder list by asking: What do they need to know? What decisions must they make? What actions must they take? This transforms a broad roster of stakeholders into a practical engagement plan.

Why it matters:

Even strong interventions can falter if key stakeholders are not aligned. Preparing for scale means bringing these actors in early, clarifying responsibilities, and ensuring they are coordinated, communicative, and moving in the same direction.

How to do it:

The Stakeholder Support Matrix is designed to help clarify the critical actors and the role each must play, with particular attention to their level of influence, the decisions they control, and the kind of engagement they will need to stay committed.

1. Outline the key actors and stakeholders involved in the scaling process

a. Building on the Stakeholder Map and Target Audiences tool created during the *Define* phase, begin by listing the actors and institutions whose decisions, resources, and actions will shape the scaling process. Scaling requires collaboration across different stakeholders groups. Some will deliver the intervention, others will provide resources, and others will guide or support the process. Scaling is more likely to succeed when each of these roles are defined from the start. Use Table 3 to identify all potential stakeholder groups.

***Note**: The examples provided in the "Why they matter" column draw from different UNICEF sectors — including child rights, protection, health and nutrition, education, WASH, and emergency response — to illustrate how scaling principles apply across diverse contexts.*

D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Stakeholder Support Matrix

Intervention: _____

List the most important actors for scaling. For each, capture their role, influence, level of support, what they need to do/know, and how and when to engage them.

Actor / Stakeholder Who matters for scaling?	**Role in scaling** Policy, delivery, advocacy, funding?	**Influence** Do they have authority to enable or block scaling (High/Med/Low)?	**Level of support / interest** High/Med/Low	**What they need to do** Decisions, actions, commitments	**What they need to know** Evidence, costs, benefits, equity implications	**Timing / entry point** When is the best moment to engage?	**Best format / channel** How to reach them?	**Engagement strategy** How to build/sustain support?

TABLE 3: COMMON KEY STAKEHOLDER GROUPS AND ACTORS TO CONSIDER

STAKEHOLDER	WHO THEY ARE	WHAT THEY DO	WHY THEY MATTER
Lead organization	The organization that will eventually own and deliver the intervention. Examples include ministries of health, non-governmental organizations, or provider networks, such as groups of clinics or hospitals.	Put the intervention into practice at scale. They deliver services, manage staff, integrate the work into budgets, and reach communities.	Their capacity and willingness to take ownership determine whether scaling is possible and sustainable. For example, when a ministry of environment adopts a community-led flood-warning system and integrates it into its national disaster-preparedness framework, the intervention becomes part of long-term climate resilience planning, rather than a project-based activity.
Resource teams	Technical partners and individuals who developed or tested the intervention. Examples include project teams, research institutions, or technical agencies.	Provide expertise on design, evidence, and adaptation. They strengthen systems and support lead organizations and stakeholder groups to adopt and deliver at scale.	They link the pilot stage to the larger system, ensuring knowledge is transferred and quality is maintained during expansion. For example, a university research team that piloted a digital learning platform for remote classrooms may support a ministry of education in adapting its content for local languages and low-connectivity schools.
UNICEF team members	UNICEF staff at headquarters, regional, and country levels who are involved in programme management, advocacy, or technical support.	Coordinate efforts, align with government priorities, and connect with other UNICEF sectors and resources.	Their influence and ability to bring partners together can speed up adoption and ensure lessons are built into wider UNICEF work. For example, UNICEF country staff may help a national civil registration authority secure funding to digitize birth registration systems, ensuring every child's identity is recorded and linked to essential services.
Funders	Donors, development banks, philanthropic organizations, or government financing bodies.	Provide financial resources, shape funding priorities, and set reporting requirements.	Lasting scale depends on reliable financing that fits both government plans and funder priorities. For example, a donor might fund the first phase of expanding early learning centres, providing the government with time to include the programme in its regular education budget.
Multipliers	Actors who can increase influence and reach. Examples include professional associations, media, regional networks, or global initiatives.	Share evidence, shape messages, advocate for policy change, and encourage an intervention's adoption.	They can spread the impact beyond the pilot by influencing both policymakers and practitioners at national and global levels. For example, when a national teachers' union endorses new child-friendly classroom practices, it encourages schools across the country to adopt the approach and strengthens policy support for safer learning environments.

STAKEHOLDER	WHO THEY ARE	WHAT THEY DO	WHY THEY MATTER
Local stakeholders	Community leaders, frontline workers, civil society groups, or traditional authorities.	Mobilize demand, adapt interventions to local norms, and maintain community trust.	Without local support, interventions risk rejection or loss once external support ends. For example, when village councils and parent groups organize community clean-up days and manage shared water points, it strengthens local ownership of safe water and sanitation services.
Other stakeholders	Partners who bring additional expertise and resources. Examples include private sector actors, universities, or ministries such as education, finance, or information technology.	Provide skills, products, or infrastructure that help with delivery and adaptation.	They add strength and resilience by bringing in resources and innovations from outside the health sector. For example, during emergencies, logistics and supply companies partnering with the government can help to ensure that relief supplies—such as safe water, learning materials, or nutrition kits—reach affected communities quickly and efficiently.

b. Start by identifying the lead organization. Begin with the brainstormed list of stakeholders and determine which one has the mandate to lead the scaling process. The lead organization is the actor with formal responsibility for the issue — such as a ministry department, government agency, or other mandated institution. The organization should also demonstrate clear demand for the intervention and have the capacity to deliver at scale. This is distinct from supportive stakeholders, who may be highly interested but lack the mandate or authority to lead.

To identify the lead organization, review Table 4, which sets out four domains for assessment: demand, capacity, timing, and context, along with strengths and constraints. Once identified, record the lead organization in Column 1 of the worksheet.

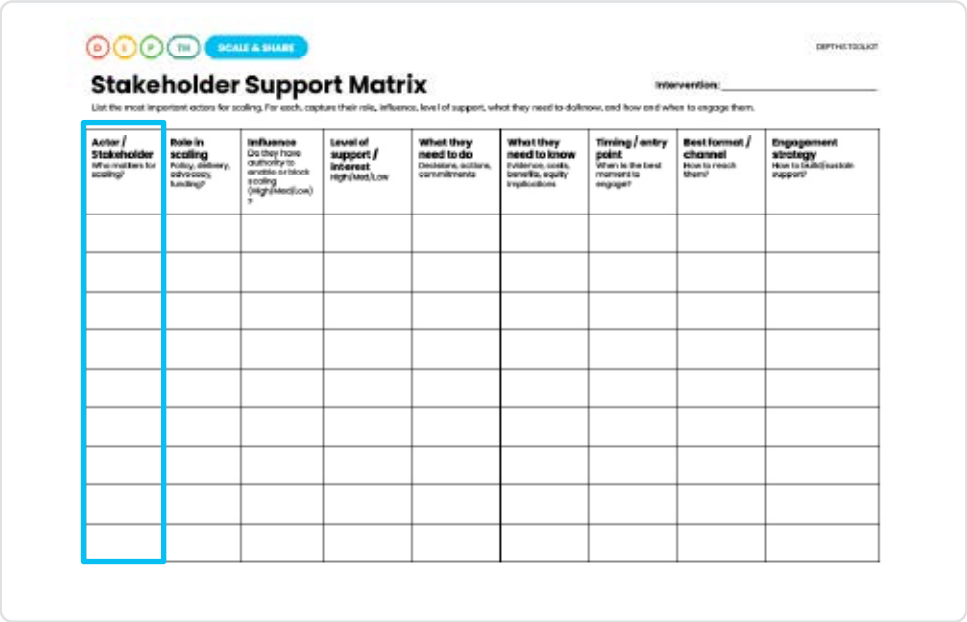
TABLE 4. GUIDE TO ASSESS THE LEAD ORGANIZATION

DOMAIN TO REVIEW	EXAMPLE: IMAGINE A COUNTRY CONSIDERING WHETHER TO SCALE UP COMMUNITY CASEWORKER HOME VISITS TO IDENTIFY AND SUPPORT GIRLS AT RISK OF CHILD MARRIAGE.
Demand: Does the intervention match the organization's priorities? Are there internal champions to push it forward? Do the benefits outweigh the costs or risks?	A ministry of social welfare may publicly commit to ending child marriage, creating strong policy demand. However, if attention and funding are focused on cash transfers or social assistance, there may be limited leadership energy to expand caseworker programmes.
Capacity: Are staff, supervisors, and supply systems ready to take on new tasks without weakening other services?	The ministry may already employ community caseworkers, but high caseloads and limited supervision could make it difficult to add new responsibilities. Additional mentoring, simplified reporting tools, or stronger links with local NGOs may be needed to sustain quality.
Timing and context: Are there opportunities such as budget cycles, new strategies, or political windows that can help? Or are there risks like elections, leadership changes, or public concerns that could slow progress?	Scaling may align with a new national strategy on child protection or a global spotlight on ending child marriage, but it could lose momentum if government restructuring shifts the programme oversight, or donor priorities change.
Strengths and constraints: Note both. Recording actions to fill these gaps creates a clear roadmap for building ownership and momentum.	The ministry may have strong coordination with local protection committees but weak digital data systems for case tracking. Identifying this gap points to the need to strengthen digital reporting tools and train staff before a national scale-up.

c. Next, identify the resource team, which supports the lead organization in moving from pilot to scale. This often includes those who designed or tested the intervention, but should also bring in capacities beyond technical expertise, such as management, financing, advocacy, and systems strengthening. To identify the resource team, review stakeholders across the relevant domains in Table 5. Record the resource team in Column 1 of the worksheet.

TABLE 5. GUIDE TO ASSESS THE RESOURCE TEAM

AREA TO REVIEW	EXAMPLE: IMAGINE A COUNTRY CONSIDERING WHETHER TO SCALE UP COMMUNITY HEALTH WORKER HOME VISITS TO INCREASE CHILDHOOD VACCINATION.
Composition: Who should be included in the resource team? Is there a mix of people with technical, managerial, and advocacy skills?	The pilot team was mainly researchers and NGO staff. For scale, the resource team now includes experts in behaviour change, supply chain management, and community engagement, alongside representatives from the Ministry of Health to ensure alignment.
Credibility and leadership: Does the team include trusted figures who can influence the lead organization and reassure communities?	A respected paediatrician known for championing immunization joins the team. Their voice gives confidence to ministry leaders and communities that home visits are credible and valuable.
Skills and experience: Does the team have the right balance of skills in supervision, financing, supply systems, monitoring, and communication? Has anyone led a scale-up before?	The NGO partner has run national training programmes, while the research team has experience in monitoring results. However, no one has led a large-scale expansion of community health worker programmes. This is flagged as a gap needing external support or mentorship.
Resources and stability: Does the team have enough staff, funding, and long-term commitment to provide ongoing support during scale?	Funding for technical support is secured only for two years. This creates a risk if scale-up requires longer accompaniment. The team records this as a gap and seeks to embed more responsibility within the Ministry of Health over time.


D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Stakeholder Support Matrix

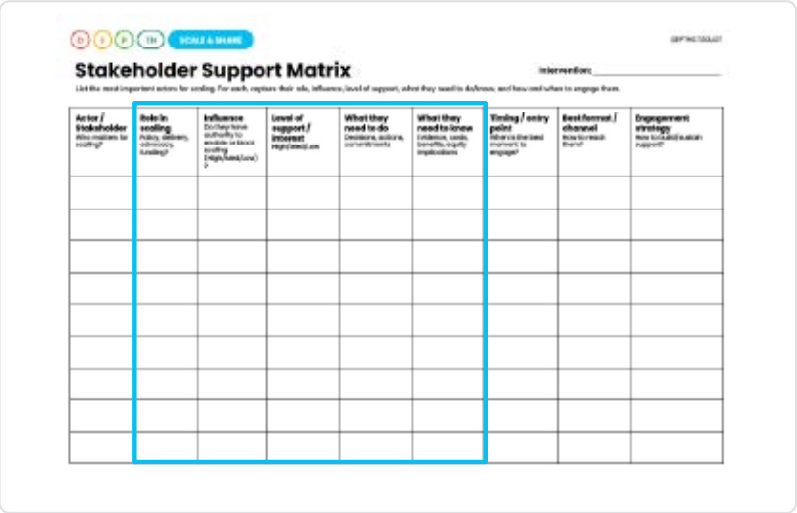
Intervention: _____

List the most important actors for scaling. For each, capture their role, influence, level of support, what they need to do/know, and how and when to engage them.

Actor / Stakeholder Who matters for scaling?	Role in scaling Policy, delivery, advocacy, funding?	Influence Do they have authority to enable or block scaling (High/Med/Low)?	Level of support / interest High/Med/Low	What they need to do Decisions, actions, commitments	What they need to know Evidence, costs, benefits, equity implications	Timing / entry point When is the best moment to engage?	Best format / channel How to reach them?	Engagement strategy How to build/sustain support?

Outside of the lead organization and resource team, scaling also relies on a wider set of stakeholders who influence acceptance, resources, and long-term support. Successful scale depends on their alignment and the effective flow of knowledge and resources between them. Identify the key actors needed for scale, and record all stakeholder groups or actors in Column 1 of the worksheet.

2. Strategize roles, influence, and support



a. **Define the role** in scaling that each stakeholder group plays in moving an intervention from a small pilot to wider delivery. Roles can include:

- **Policymaking**: setting standards or adding an intervention to national guidelines
- **Financing**: securing budgets or donor support
- **Delivery**: managing staff and services
- **Supervision**: maintaining quality and accountability
- **Advocacy**: mobilizing champions or shaping public opinion

Roles of stakeholder groups can change over time. A project team may lead during the pilot, but later shift into a support role once a ministry or provider network takes ownership. Clarifying roles allows teams to coordinate efforts, manage handovers, and ensure that scaling responsibilities are clear, achievable, and sustainable.

b. Next, capture the degree of **influence** each stakeholder group has to enable or block progress. Influence can be formal, such as approving budgets or passing regulations, or informal, such as trusted community leaders shaping public views. Mapping influence helps to highlight who the decision makers are, where alliances are needed, and where advocacy should be directed. For instance, a professional association might not control funds but can sway health workers' acceptance.

c. Assess the level of **support/interest** from each stakeholder group for the intervention. High support often comes when the intervention matches existing priorities or personal commitment. Low support may come from competing agendas or doubts about effectiveness. Recognizing these drivers helps to anticipate resistance and tailor engagement. For example, a district health office may be enthusiastic if the programme reduces their workload, but less interested if it adds extra reporting requirements.

Clarify what each stakeholder group **needs to do**. This avoids vague promises and links groups and people to concrete actions, such as approving policy changes, allocating resources, or delivering services. For instance, a ministry may be responsible for printing materials at national level, while local managers ensure they reach clinics on time.

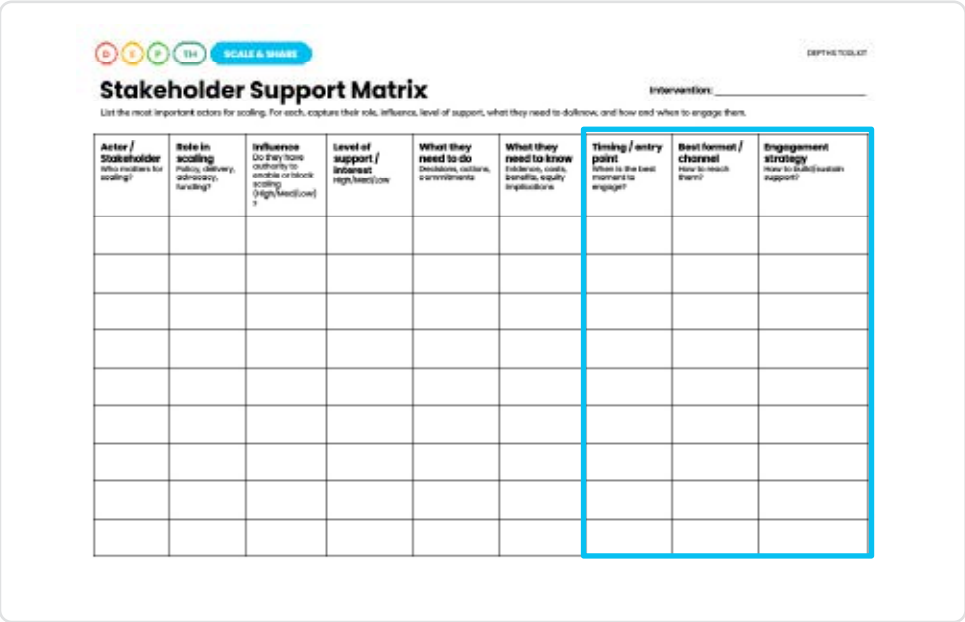
Identify what each stakeholder group **needs to know** in terms of evidence or information required to take action. Policymakers may need data on costs and equity, funders may need clear reporting frameworks, and community leaders may want proof that the intervention improves local wellbeing. Tailoring evidence to these needs ensures that information is both useful and motivating.

3. Plan stakeholder engagement

a. Determine the **timing and entry points** of engaging with each stakeholder group. Some groups work on fixed cycles such as budget approvals or policy reviews. Others may be most receptive during pilot demonstrations or community consultations. Planning around these windows prevents missed opportunities. For example, introducing results just before an annual budget cycle can increase the chance of funding.

b. Choose the best **format/channel** to communicate with stakeholder groups. For example, senior government leaders may prefer short policy briefs or direct meetings. Implementers may engage better through workshops, while communities may respond to radio messages or local events. Choosing the right format increases clarity, credibility, and uptake.

c. Plan the best **engagement strategy** for each stakeholder group, outlining how to build and maintain support over time. This can involve targeted advocacy, regular relationship management, or nurturing champions within institutions. The goal is to turn early interest into long-term commitment. For example, assigning a respected district officer as a champion can maintain momentum even when leadership changes.



D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Stakeholder Support Matrix

Intervention: _____

List the most important actors for scaling. For each, capture their role, influence, level of support, what they need to do/know, and how and when to engage them.

Actor / Stakeholder Who matters for scaling?	Role in scaling Policy, delivery, advocacy, funding?	Influence Do they have authority to enable or block scaling (High/Med/Low)?	Level of support / interest High/Med/Low	What they need to do Decisions, actions, commitments	What they need to know Evidence, costs, benefits, equity implications	Timing / entry point When is the best moment to engage?	Best format / channel How to reach them?	Engagement strategy How to build/sustain support?

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The tools mentioned in this step of the Scale phase were not used by the original project team. This case study is a recreated example based on real project data and context.

Outlining the key stakeholders

The team began by revisiting the initial Stakeholder Map they had created at the beginning of the project, from the Define phase. They analysed and then built the list with new stakeholders they considered important for the scaling process. For example, some key stakeholder groups mapped were the Ministry of Public Health's Expanded Programme on immunization, primary health care centres (PHCCs), potential funders such as Gavi and the European Union, paediatric associations and local media, and local stakeholders such as community leaders, health workers, and NGOs.

Identifying the lead organization

The Ministry of Public Health was identified as the lead organization. This choice was based on its mandate and capacity to oversee national immunization services. The team assessed the ministry across four domains:

- Demand: Vaccination was a top health priority, and champions within the EPI unit signalled strong interest in the postcards.
- Capacity: Supply and distribution systems were relatively robust, but staff would need training to use the postcards consistently without extra supervision.
- Timing and context: Budget discussions were underway, creating an opportunity to propose a new budget line for printing. At the same time, leadership changes in the ministry posed some uncertainty, making advocacy at multiple levels essential.
- Strengths and constraints: The Ministry had strong procurement systems but weak community outreach. The team noted this, and noted that partnerships with NGOs could help to fill this gap.

Identifying the resource team

The resource team included UNICEF staff and the researchers who had led the trial. Together they brought technical evidence, credibility, and experience. The team assessed themselves using the four domains:

- Composition: While researchers brought technical knowledge, the team added UNICEF staff with experience in government engagement and supply chain management.
- Credibility and leadership: A respected paediatrician was invited as an advisor to increase trust among ministry leaders and communities.
- Skills and experience: The team had expertise in evaluation and communication, but limited experience in national scale-up. They flagged this as a gap and sought mentorship from UNICEF colleagues who had worked on other scaled health interventions.
- Resources and stability: External funding was secured for two years, enough to support initial expansion but not long-term scale. This made government budget allocation a critical goal.

Defining other stakeholders

Beyond the Ministry and the resource team, scaling depended on wider stakeholder groups. UNICEF staff at the country office provided policy alignment and convening power. Funders offered financial support and legitimacy. Multipliers such as professional associations helped to spread visibility and endorsement. Local stakeholders, community leaders, NGOs, and frontline staff ensured the intervention remained trusted and relevant in daily practice. Together, these groups formed an ecosystem where each stakeholder group's role was distinct but interconnected.

Determining roles, influence, and support

The Stakeholder Support Matrix helped to clarify each stakeholder group's role. The Ministry of Public Health held high authority and medium-to-high support, with responsibilities ranging from issuing circulars to financing. Funders had high influence through budget priorities. UNICEF played a bridging role, coordinating between

technical experts and the government. Multipliers and local stakeholders had lower formal authority but high informal influence, shaping public trust and acceptance.

What they needed to do and know

The team defined concrete responsibilities for each stakeholder. The Ministry needed to approve a budget line, update guidelines, and supervise delivery. Funders needed to commit short-term financing. PHCC staff had to distribute and explain postcards to caregivers. Local stakeholders needed to encourage families to keep and use the cards. To act, each group required tailored information: policymakers needed cost and equity data, funders needed evidence of effectiveness, and communities needed reassurance that the intervention was simple, safe, and useful.

Timing, channels, and engagement strategies

Engagement was planned around natural entry points. For the Ministry, the optimal moment was the annual budget cycle and the next meeting of the National Immunization Technical Advisory Group. For funders, the key window was the donor coordination forum. For local stakeholder groups, entry points were routine community meetings and PHCC consultations.

Communication formats were tailored: a short policy brief for the ministry, a costing slide deck for funders, and pictorial materials for communities. Strategies also varied: high-level advocacy with government, relationship management with funders, and participatory engagement with community leaders.

Putting it all together

By using the Stakeholder Support Matrix, the team turned a long list of potential stakeholder groups into a practical plan for engagement. Each stakeholder was matched with roles, responsibilities, and entry points. This ensured that when horizontal and vertical scaling plans were later developed, they were grounded in real commitments and aligned with the institutions that would sustain the intervention.

D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Stakeholder Support Matrix

Intervention: Appointment reminder card

List the most important actors for scaling. For each, capture their role, influence, level of support, what they need to do/know, and how and when to engage them.

Actor	Role in scaling	Influence	Level of support	What they need to do	What they need to know	Timing / entry point	Best format / channel	Engagement strategy
Ministry of Public Health (MoPH)	Policy, financing, supervision	High	High	Endorse reminder postcards as part of national immunization strategy, include in guidelines, allocate budget for printing and distribution	Evidence of improved vaccination uptake, affordability data (low per-child cost), potential for equity gains in underserved areas	Annual immunization programme review, budget cycle	Policy briefs, technical presentations, high-level meetings	Position postcards as a cost-effective, scalable intervention that strengthens routine immunisation. Link to national child health priorities.
Primary Health Care Centres (PHCCs)	Delivery, supervision	Medium	Medium to High	Distribute postcards, train staff on filling and explaining them, track use through supervision	Evidence on simplicity of process, minimal additional workload, improved caregiver return rates	During staff training sessions and supervision cycles	Practical training sessions, illustrated job aids	Emphasise how postcards reduce missed appointments and ease follow-up work. Provide supportive supervision to reinforce use.
Community health workers (CHWs)	Delivery, advocacy	Low to Medium	Medium	Encourage caregivers to use and keep postcards, answer questions, follow up in the community	Benefits for caregivers (fewer missed vaccinations, healthier children), evidence of ease of use	Community meetings, monthly outreach cycles	Orientation sessions, simple pictorial materials	Position postcards as a tool that strengthens trust with families. Reinforce role as trusted messengers.
Caregivers / communities	End users, advocacy	Low individually, High collectively (through uptake and word of mouth)	Variable (High where trust in vaccination is strong, lower where hesitancy exists)	Bring postcards to visits, follow reminders, encourage peers	Evidence that postcards help protect children, stories of other caregivers benefiting, reassurance on confidentiality	Community gatherings, vaccination sessions	Posters in clinics, community radio, peer mothers' groups	Build demand through visible use of postcards, testimonials, and community endorsement.
UNICEF Lebanon	Technical support, advocacy, funding leverage	High	High	Advocate with government, align with broader child health strategies, provide initial funding and technical assistance	Evidence of effectiveness from the RCT, costs, and relevance to UNICEF's child health agenda	Policy discussions with MoPH, donor meetings	Policy briefs, technical reports, joint MoPH-UNICEF presentations	Use UNICEF credibility to reinforce government ownership, ensure alignment with global immunisation initiatives.
Donors (e.g., Govt. EU, WHO, local donors)	Funding, advocacy	High	Medium to High	Provide financing for early scale-up, integrate into grants or projects	Cost-effectiveness data, sustainability pathways, potential impact at national level	Donor funding cycles, proposal windows	Funding proposals, results briefs, donor roundtables	Frame postcards as a "quick win" intervention with strong evidence and low cost, while highlighting integration into MoPH systems for sustainability.
Local NGOs / CSOs	Advocacy, community engagement	Medium	Medium	Help raise awareness, integrate postcards into outreach, provide feedback from the ground	Local impact stories, evidence of caregiver acceptance, role in reducing dropouts	During project rollouts, community mobilisation activities	Community meetings, joint workshops with MoPH/UNICEF	Strengthen trust and local ownership by involving them in adaptation and feedback loops.

STEP 3:
Plan for Scale

Associated Tools:

- Horizontal Scaling Plan
- Vertical Scaling Plan

In this step:

The final step is to develop a scaling plan for the pathway identified in Step 1. Scaling requires deliberate choices about whether to grow outwardly to new settings, districts, or communities (horizontal scaling) or embedding the intervention into policies, budgets, and institutions (vertical scaling). Both can support long-term sustainability, but each has distinct considerations.

Horizontal scaling builds sustainability by showing that an intervention works across diverse contexts, and by spreading demand and ownership amongst diverse communities. The challenge is to preserve the core elements that make it effective while adapting to new conditions.

Vertical scaling secures resources and long-term commitment by institutionalizing the intervention within government systems. Its success, however, depends on political will and institutional capacity.

Whichever pathway is chosen, success depends on protecting the intervention's core elements and building the partnerships, systems, and resources needed for them to endure at scale.

Why it matters:

Evidence from a pilot may show that an intervention works, but without a strategy for growth it can remain an isolated success. Scaling is shaped not only by design, but also by politics, resources, and institutions. A clear plan makes these forces visible, helps teams to anticipate challenges, and guides trade-offs. It clarifies the roles of different actors and ensures systems and resources are in place to carry the impact forward.

How to do it:

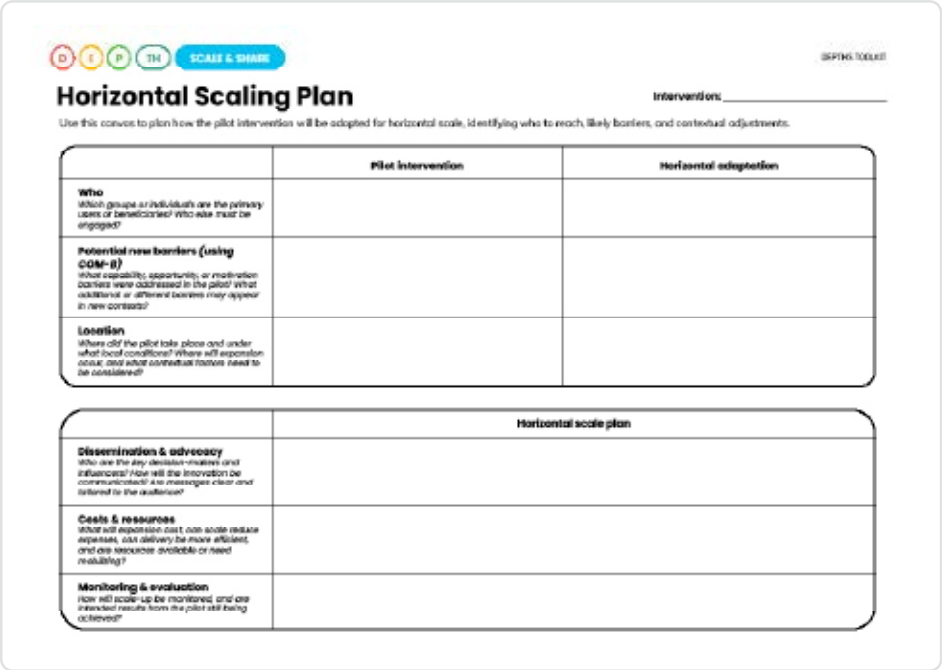
The tools in this step help to identify the core elements that must be preserved, the adaptations required, and the new barriers or demands that may emerge. The result is a concrete roadmap for scaling.

1. Horizontal Scaling: Expanding reach

Using part I of the Horizontal Scaling Plan, compare the realities of the pilot with the adaptations needed for expansion across key dimensions: where the intervention will be implemented, who it is intended to reach, and how behavioural barriers may shift during scale.

a. Starting with "**Who**" in the top box of the worksheet, record the primary users or beneficiaries from the pilot, along with the actors who supported delivery. Then consider expansion: as the intervention reaches new geographies or populations, will these same groups remain central, or will additional ones need to be engaged?

- **Expanding to new geographies:** A pilot in one district may have relied on a small group of dedicated teachers and community youth volunteers. Scaling to several districts may require mobilizing additional stakeholders such as faith leaders, local councils, or parent associations who hold influence in those new contexts.
- **Expanding to new population groups:** When scaled, a pilot designed for out-of-school adolescents may also need to engage parents, teachers, and community mentors to support enrolment and sustained participation.



D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Horizontal Scaling Plan

Intervention: _____

Use this canvas to plan how the pilot intervention will be adapted for horizontal scale, identifying who to reach, likely barriers, and contextual adjustments.

	Pilot Intervention	Horizontal adaptation
Who *Which groups or individuals are the primary users or beneficiaries? Who else must be engaged?*		
Potential new barriers (using COM-B) *What capability, opportunity, or motivation barriers were addressed in the pilot? What additional or different barriers may appear in new contexts?*		
Location *Where did the pilot take place and under what local conditions? Where will expansion occur, and what contextual factors need to be considered?*		

	Horizontal scale plan
Dissemination & advocacy *Who are the key decision-makers and influencers? How will the innovation be communicated? Are messages clear and tailored to the audience?*	
Costs & resources *What will expansion cost, can scale reduce expenses, can delivery be more efficient, and are resources available or need re-building?*	
Monitoring & evaluation *How will scale-up be monitored, and are intended results from the pilot still being achieved?*	

In both cases — whether expanding to new geographies or to new population groups — horizontal scaling requires clarifying how the who changes from pilot to scale. This includes not only those directly reached by the intervention, but also the intermediaries and influencers whose involvement is essential to sustain and expand impact. Mapping these shifts sets the stage for the next step: identifying how new groups may introduce new barriers that must be anticipated for the intervention to succeed at scale.

b. The next step is to examine how the **behavioural drivers** of these new groups may differ from those in the pilot. Applying the COM-B framework provides a structured way to identify potential barriers to capability, opportunity, or motivation at scale, and to compare them with the barriers already addressed during the pilot.

- **Expanding to new geographies:** In one setting, reminders may reduce a capability barrier by helping people to remember contraceptive refills or follow-up visits. But when the same intervention is introduced in rural areas, new barriers can appear. Longer distances to social service offices may prevent families from accessing legal or child-protection support (opportunity barrier), and in some areas caseworkers may have limited training to identify and respond to child-safety concerns (capability barrier).
- **Expanding to new population groups:** A school attendance reminder system originally designed for parents may need to be adapted for older students themselves. This can introduce new barriers — for example, some adolescents may see attendance reminders as unnecessary or intrusive (motivation barrier), or may not know how to update or respond to messages in the platform (capability barrier).

  By comparing the behavioural barriers addressed in the pilot with those likely to emerge in new geographies or population groups, teams can pinpoint the adaptations needed to keep the intervention effective.

c. Scaling requires carefully comparing the **conditions** of the pilot with those in the locations where scale will occur. Geography, infrastructure, service delivery systems, and local dynamics can all help or hinder effectiveness. What factors supported success in the pilot, and what will differ in the new context? Documenting these differences prevents the assumption that what worked in one place will automatically succeed elsewhere.

- **Expanding to new geographies:** A programme that thrived in an urban district with reliable transport, well-equipped schools, and steady teacher staffing may struggle in rural areas where schools are far apart, transport is costly and irregular, and resources arrive inconsistently. These differences can make it harder for children to attend regularly and for teachers to deliver lessons effectively.
- **Expanding to new populations**: A pilot tested in one linguistic or cultural setting may need significant adaptation when introduced into areas with different languages, social norms, or behaviours. For example, messages designed for communities where mothers make health decisions may not resonate in places where fathers or elders play a stronger role, or where translation into local languages changes the clarity or tone of the message.

The second part of the Horizontal Scaling Plan moves to planning the additional elements needed for success at scale. Teams must consider how to communicate and advocate for the intervention, how to secure and allocate resources, and how to monitor and evaluate progress so that results are sustained as the intervention's reach expands.

d. Begin with **dissemination and advocacy**. Scaling to new regions or populations depends on building support from the people and institutions who can enable or block progress. This section draws directly on the Stakeholder Support Map from Step 2 and translates that analysis into a clear plan for action.

Use the worksheet to produce a single, coherent, and accessible paragraph that answers three core questions:

- Who are the key decision-makers and influencers?
- How will the intervention be communicated to them?
- Are the messages tailored to their priorities?

e. Next, reflect on **costs and resources**. Expansion requires deliberate decisions about how costs will shift across settings and whether resources can be sustainably mobilized at scale. This part of the Horizontal Scaling Plan focuses on assessing both affordability and long-term sustainability, anticipating uneven cost patterns across contexts rather than assuming uniformity.

Horizontal scaling often reveals hidden variations. For example, rural areas may demand additional transport, supervision, or incentives to reach dispersed populations, while urban areas may require greater investment in communication channels and community outreach. Some helpful guiding questions to better plan for costs and resources needed during horizontal scaling are:

- **What will expansion cost across different settings?** Estimate the financial requirements for training, supplies, transport, supervision, and communication in new geographies or populations. Consider where costs may rise and where efficiencies of scale may offset them. For example, rural districts may require extra fuel and allowances for outreach teams, while in dense urban areas, it's common to spend more on community outreach.
- **What existing resources can be leveraged?** Review which resources supported the pilot, such as staff time, infrastructure, or volunteer contributions, and assess whether these can stretch to support scale. Identify gaps that must be filled. For example, school committees and local water-user groups can organize hygiene activities during community meetings, and existing maintenance logs or mobile apps can help to track when repairs are due.
- **How will new resources be mobilized?** Determine whether expansion will be funded through district or national budgets, donor partnerships, or integration into sector-wide financing mechanisms. Plan strategies to secure and sustain resources so that costs do not become a barrier to scale. For example, UNICEF can help the Ministry of Education include training materials and teacher stipends in the annual education budget and partner with local radio stations to provide learning content at a reduced cost.

f. Finally, plan for **monitoring and evaluation (M&E)**. As interventions expand into new geographies or populations, monitoring systems must balance comparability with adaptability: indicators should be consistent enough to track performance across sites, yet flexible enough to capture local adaptations. For example, a national child helpline might track a core indicator such as "number of calls responded to within 24 hours," while allowing country-specific measures — like "calls received in local languages" or "percentage of referrals successfully completed" — to reflect context-specific priorities.

M&E should also assess whether scale is reaching all intended groups, whether outcomes are sustained across diverse populations, and whether the core behavioural mechanisms that drove the pilot's success remain intact. Comparisons between pilot and scale sites are especially important for spotting where results weaken and why.

For example, if a pilot reading campaign improved attendance and literacy outcomes in urban schools, M&E at scale should include measures of lesson delivery and student engagement, to ensure that children in rural or resource-constrained schools are equally reached and benefiting.

Below are some helpful guiding questions to consider when building an M&E plan.

- **How will success be measured?** M&E plans should be designed to measure more than reach. Data should be collected to capture whether the intervention is being delivered as intended (process), whether it continues to generate the desired changes (outcomes), and whether it contributes to longer-term shifts (impacts). Select a small set of indicators that balance feasibility (data that can realistically be collected across multiple sites) with meaningfulness (data that confirms the intervention is generating impact as intended, and that behavioural mechanisms are holding).
- **What systems and tools will be used?** Sustainability depends on embedding monitoring into existing structures wherever possible. Explore how to integrate new indicators into government systems or existing service statistics. Where routine systems are weak, consider lighter supplemental tools, like simple reporting forms or mobile-based tracking, that can provide timely data without overburdening staff.

- **What additional assessments are needed?** Routine monitoring rarely captures the full picture. Rapid qualitative studies, focus group discussions, or community feedback tools can help to explain why outcomes differ across sites, while special evaluations may be needed to test outcomes or impacts more rigorously. These complementary assessments ensure that scale-up is not only tracked but also understood.
- **How will findings inform strategy?** Data is only valuable if it is used. Build in regular review mechanisms (monthly dashboards, quarterly reflection meetings, or joint reviews with stakeholders) to ensure that results feed directly into decision-making. Findings should guide any adjustments to implementation, refine delivery strategies, and hold partners accountable for supporting scale.

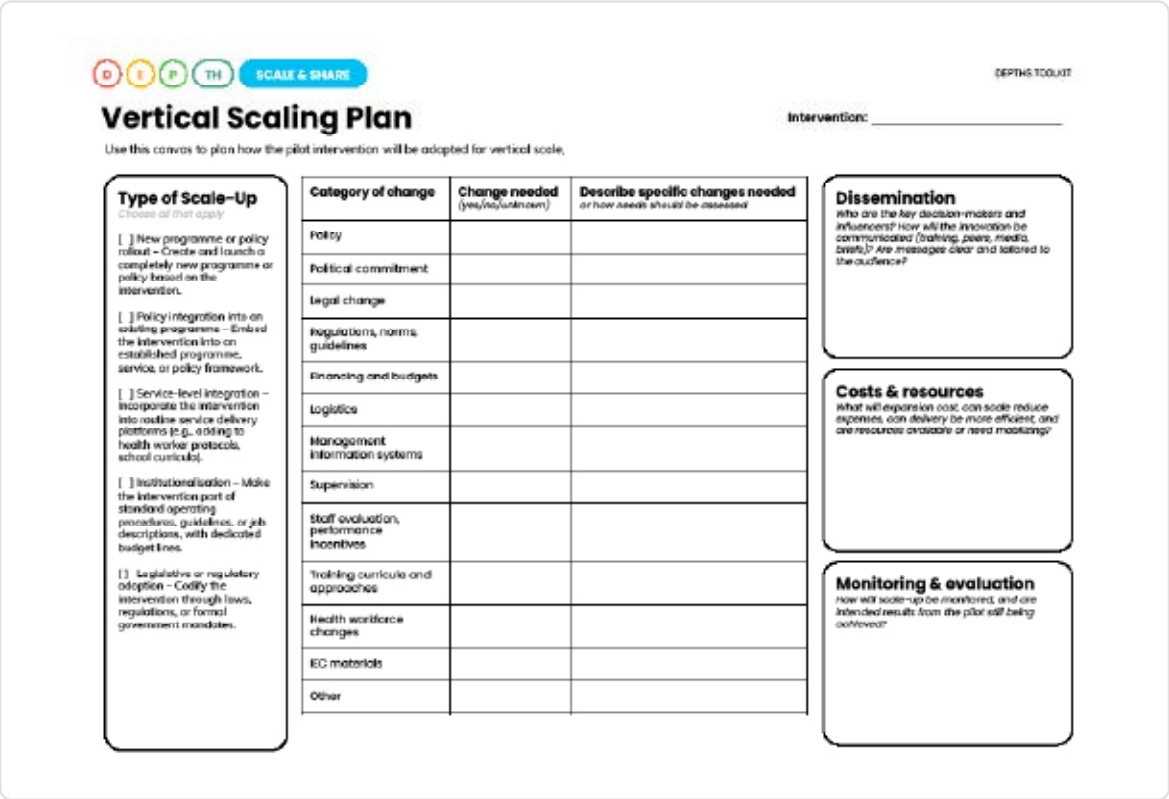
2. Vertical scaling: Institutionalization and systems change

While horizontal scaling is a more common pathway, vertical scaling is just as important. When an activity is built into systems, policies, and budgets, it becomes part of routine delivery and delivers lasting results. Vertical scaling depends on alignment with policy makers, funders, civil servants, and the parts of government that can formalize and sustain actions. In short, success depends on the extent to which activities can fit within existing institutions and systems.

For example, a school nutrition activity might work well in a pilot, but it will only last if it is written into teacher training, included in education budgets, and monitored through school inspections.

The Vertical Scaling Plan helps to identify categories of change, specify the necessary adjustments, and outline how the change will be communicated, funded, and monitored.

a. The first step is to clarify the type of vertical scale-up in the left-hand column of the worksheet. Ask: What form of scale-up is the goal? The table below summarizes common types of vertical scale-up and offers guidance on when each type can be most useful.



D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Vertical Scaling Plan

Intervention: _____

Use this canvas to plan how the pilot intervention will be adopted for vertical scale.

Type of Scale-Up

Choose all that apply

[ ] New programme or policy rollout – Create and launch a completely new programme or policy based on the intervention.

[ ] Policy integration into an existing programme – Embed the intervention into an established programme, service, or policy framework.

[ ] Service-level integration – Incorporate the intervention into routine service delivery platforms (e.g., adding to health worker protocols, school curricula).

[ ] Institutionalisation – Make the intervention part of standard operating procedures, guidelines, or job descriptions, with dedicated budget lines.

[ ] Legislative or regulatory adoption – Codify the intervention through laws, regulations, or formal government mandates.

Category of change	Change needed (yes/no/unknown)	Describe specific changes needed *or how needs should be assessed*
Policy		
Political commitment		
Legal change		
Regulations, norms, guidelines		
Financing and budgets		
Logistics		
Management information systems		
Supervision		
Staff evaluation, performance incentives		
Training curricula and approaches		
Health workforce changes		
IEC materials		
Other		

Dissemination

Who are the key decision-makers and influencers? How will the innovation be communicated (training, peers, media, briefs)? Are messages clear and tailored to the audience?

Costs & resources

What will expansion cost, can scale reduce expenses, can delivery be more efficient, and are resources available or need mobilizing?

Monitoring & evaluation

How will scale-up be monitored, and are intended results from the pilot still being achieved?

TABLE 7. TYPES OF VERTICAL SCALE-UP

TYPE OF SCALE-UP	MOST USEFUL WHEN	EXAMPLE
New programme or policy rollout: Create and launch a brand-new national or sub-national programme built around the intervention.	There's no obvious home for the intervention; high visibility and coordinated roll-out are needed.	The Ministry of Health launches a National Follow-Up Calling Programme, mandating a weekly two-hour calling block in all clinics, funded by the government.
Integration into existing policies or programmes: Embed the intervention within an established programme so it becomes a standard component.	An existing programme can absorb the work with minor adjustments.	The national immunization programme updates its policy to include weekly defaulter calls, adds it to annual plans, and aligns it with supervision and reporting.
Service-level integration: Incorporate the intervention into routine workflows and supervision.	The main change is frontline practice rather than high-level policy; quick adoption is feasible.	Clinics add a two hour weekly calling slot to timetables, nurses use a 60 second script, and supervisors check a one page call log during regular visits.
Institutionalization (SOPs, roles, budgets): Make the intervention part of standard procedures, job descriptions, supervision tools, indicators, and give it a dedicated budget line.	Long-term durability and funding protection are needed; shifting from project to "business as usual."	Nurse job descriptions include calling caregivers of missed appointments, supervision forms add a tick-box, a small monthly airtime line is in clinic budgets, and a simple calling indicator is included in routine reports.
Legislative or regulatory adoption: Codify the intervention through regulations or laws to set authority, privacy standards, and minimum practice.	Legal clarity is needed (e.g., privacy/consent), leadership turnover is likely, or standards must be protected across administrations.	A health regulation authorizes brief clinic calls to caregivers using an approved privacy script, sets data-handling rules, and requires basic call records.

b. The next step is to check which parts of the system must change so the activity can move from a pilot into a policy or routine delivery. The middle column of the Vertical Scaling Plan lists areas to review: policy, political commitment, financing, supervision, training, and information systems. For each area, note if a change is needed and describe the specific adjustment. This comparison shows where institutional change is required to move beyond short term projects.

- **Policy**: Does scaling require a new policy directive or inclusion in an existing policy framework? For example, adding community caseworker roles and reporting protocols to the national child protection policy so that outreach becomes part of routine government service delivery.
- **Political commitment**: Are champions at higher levels of government needed to support institutionalization? Political leadership may be critical to secure approval or drive adoption.
- **Financing and budgets**: Will new budget lines be needed, or can financing be absorbed into existing sector allocations? For instance, including teacher training

costs for a new reading initiative within the national education budget rather than relying on external project grants.

- **Supervision and performance management:** Does institutionalization demand new supervisory structures, evaluation criteria, or reporting lines? For example, ensuring district managers include the intervention in their regular oversight.
- **Training and curricula**: Are pre-service or in-service training materials required to institutionalize practices for health workers, teachers, or other frontline staff?
- **Information systems**: Should new indicators be added to administrative datasets, school records, or other reporting platforms to ensure sustainability and accountability?

c. Next, reflect on **dissemination needs**. Vertical scale depends on visibility, legitimacy, and shared ownership among institutions that set policy and control finance. At this stage, the focus shifts from sharing results to advocating for system change, using channels that influence decision makers, budget holders, and institutional leaders.

Advocacy can include short policy briefs with clear recommendations, meetings with key ministries and partners, identifying internal champions, and site visits where leaders see the activity in practice. Input into national or subnational processes, such as budget reviews and sector plans, is often the moment when routine adoption can be secured.

A core aim is broad ownership. Present the activity as part of routine delivery, not an add-on. Tailor messages for senior officials, programme managers, frontline staff, and community representatives for each to understand their role in sustaining the work.

By matching formal and informal channels to upcoming policy moments and institutional priorities, the activity can be seen as credible and essential for routine practice.

This section of the worksheet should distil insights from the Scaling Support Matrix in Step 2 into a focused plan. It should identify priority actors, outline how to reach them, and confirm that messages are simple, tailored, and aligned with institutional priorities.

d. Vertical scale requires dedicated **resources**. Pilots often rely on external or short term funds, but institutionalization depends on fitting costs into government and partner budgets. The key task is to show how ongoing costs for training, supervision, logistics, and supplies will sit within existing budgets for health, education, or social welfare. The aim is to move from temporary funding to stable investment, so the activity is not exposed to any project cycles or donor changes.

This section should produce one clear paragraph that answers the following four questions:

- What will vertical expansion cost?
- Can scale reduce costs or make delivery more efficient?
- Which resources are already available?
- Which new resources are needed, and through which budgets or partnerships will they be raised?

e. Finally, plan for **monitoring and evaluation (M&E)**. Vertical scale should track both results and progress in institutionalization. Pilots focus on behaviour change and service use. At scale, monitoring must also show whether the activity is being built into systems, policies, and budgets, and whether it continues to deliver results once embedded. For example, track whether dedicated budget lines exist for printing reminders, and whether supervisors record that reminders were given during routine visits.

Begin by asking: How will vertical scaling be monitored and evaluated, and what indicators are appropriate? This may involve:

- **Tracking institutionalization milestones:** New policies, budget lines, updates to training, or simple indicators added to routine records.
- **Monitoring outcomes at scale:** Ensuring that the behavioural results demonstrated in the pilot are still being achieved.

- **Using existing service statistics:** Integrating indicators into routine reporting systems, so that monitoring is sustainable and aligned with sector priorities.
- **Supplementing with additional studies:** Rapid qualitative research to identify barriers as the activity becomes routine, and simple checks to see if it is being delivered as planned and still achieving results.
- **Realizing course corrections:** Regular reviews that use findings to adjust plans when political, financial, or system bottlenecks appear.

Combining institutional indicators with outcome data helps to confirm that an activity is anchored in systems that can sustain it over time. For example, if a school-based handwashing initiative improved attendance and hygiene outcomes during the pilot, national scale-up should also track whether the Ministry of Education has included hygiene promotion in school supervision checklists, teacher training plans, and annual budgets — signs that the practice is becoming part of routine delivery.

This section of the worksheet should produce one clear paragraph that sets out the monitoring approach. It should name the key institutionalization indicators to measure, state how outcomes will be checked at scale, and explain how findings will feed into decisions, so strategies can be adjusted in real time. The aim is a concise plan that shows both how institutionalization will be measured, and how effectiveness will be maintained.

The Vertical Scale Plan helps ministries, partners, and funders align on priorities, set a realistic order of actions, and track progress over time. It also clarifies who is responsible for policy changes, budgets, and monitoring, so commitments turn into routine practice. The plan should be treated as a living document and reviewed at set intervals, with clear triggers for course correction if risks or delays appear.

The tools in this phase help to test readiness, ground scaling efforts in evidence and alignment, and set clear paths for horizontal adaptation and vertical institutionalization — turning promising pilot results into lasting change.

CASE STUDY:

Increasing childhood vaccination uptake in Lebanon

The tools mentioned in this step of the Scale phase were not used by the original project team.
This case study is a recreated example based on real project data and context.

Choosing the pathways

With intervention scale readiness established in Step 1 and the ground prepared in Step 2, the project team faced the core choice of how to scale: expand across settings (horizontal) while also embedding within systems (vertical). The Stakeholder Support Map suggested both routes were viable: district managers and community actors were ready for expansion, and the Ministry of Health's Expanded Programme on Immunization signalled intent to institutionalize the approach, pending a short phased roll-out.

The team planned a two track scale: a controlled horizontal expansion to three new regions with different geographic and population characteristics, alongside a vertical integration process to embed the intervention in national policy, budgets, and supervision systems.

Horizontal Scaling Plan — Part I:

Pilot to scale comparison. The team confirmed who remained central: caregivers as primary users, outreach workers and supervisors as implementers, and EPI/UNICEF as enablers. They listed anticipated COM-B shifts at scale:

- **Capability**: A higher share of low-literacy caregivers in rural areas, requiring pictorial and multi-language card versions.
- **Opportunity**: The risk of inconsistent distribution in districts without established outreach channels, and reliance on standard (not project-specific) supervision.
- **Motivation**: Novelty could fade if the card design remained unchanged.

They also noted contextual differences between the pilot and new sites. The original pilot operated in one urban

D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Horizontal Scaling Plan

Intervention: Appointment reminder card

Use this canvas to plan how the pilot intervention will be adapted for horizontal scale, identifying who to reach, likely barriers, and contextual adjustments.

	Pilot intervention	Horizontal adaptation
Who *Which groups or individuals are the primary users or beneficiaries? Who else must be engaged?*	Primary users/beneficiaries: Caregivers of un- or under-vaccinated children, mainly in vulnerable households.	Other engaged actors: Outreach workers delivering postcards, supervisors overseeing fidelity, Ministry of Health approving design, UNICEF providing technical support.
Potential new barriers (using COM-B) *What capability, opportunity or motivation barriers were addressed in the pilot? What additional or different barriers may appear in new contexts?*	• Capability: Caregivers lacked planning capability (forgetting dates) • Motivation: Low salience • Opportunity: Weak prompts to attend clinics.	• Capability: higher-literacy groups may ignore visuals, while low-literacy groups may need more pictorial adaptation. • Opportunity: Distribution may be inconsistent if supply chain is weak. • Motivation: Novelty may fade if postcards are not refreshed; competing priorities in households may reduce motivation to act.
Location *Where did the pilot take place and under what local conditions? Where will expansion occur, and what contextual factors need to be considered?*	Implemented in select low-resource communities with strong outreach presence, during a short-term vaccination campaign.	Planned for routine immunisation in urban, peri-urban, and rural areas. Contextual factors include varying literacy rates, mobile/migrant populations, and different levels of trust in government health services.

	Horizontal scale plan	
Dissemination & advocacy *Who are the key decision-makers and influencers? How will the innovation be communicated? Are messages clear and tailored to the audience?*	Key decision-makers and influencers: Ministry of Health leadership, donor agencies, district health directors, outreach supervisors. Tailoring: Use multi-language and pictorial formats for caregivers; emphasise cost-effectiveness and equity for policymaker/donors.	Communication approach: • Policymakers: policy briefs with cost-benefit results. • Donors: slide decks showing equity and cost-effectiveness. • Outreach workers: simple training guide cards. • Communities: human-interest stories and local demonstrations.
Costs & resources *What will expansion cost, can scale reduce expenses, can delivery be more efficient, and are resources available?*	Pilot costs: $0.20 per postcard, less than $10,000 total for 10,000 households. Scaling estimate: Printing 500,000 postcards annually (~ $100,000 - with potential bulk printing discounts-) Resources needed: Funding for printing and replenishment, integration with vaccine supply logistics, and small training modules. Efficiency: Economies of scale likely, especially if bundled with vaccine shipments and procured nationally.	
Monitoring & evaluation *How will scale-up be monitored, and are intended results from the pilot still being achieved?*	Monitoring methods: Supervisors check postcard distribution during outreach visits. Routine immunisation registers track uptake and completion. Household caregiver surveys capture recall and usefulness. Customised templates for procurement and logistics.	Success criteria: (1) More than 5 percentage point increase in timely vaccine completion across diverse contexts; (2) Fidelity of 80% (postcards correctly filled and delivered); (3) equity gap reduction between vulnerable and non-vulnerable populations; (4) stable cost per additional vaccinated child (less than $10).

district during an intensive campaign period, whereas scale-up involved year-round routine delivery across three new regions, including remote and peri-urban areas with more limited transport and staff coverage.

Horizontal Scaling Plan — Part II: Dissemination, resources, and monitoring

The plan translated comparison into action. District managers were slated to receive a concise "why this, why now" brief, supervisors would run a short practice during the monthly meeting, and clinics would use a simple pictorial job aid to demonstrate the postcard to caregivers. Costs and resources were written directly into the plan, including a printing contract to keep the unit cost below twenty cents and an initial requirement of roughly 150,000 cards for the first three regions. Distribution was riding on existing vaccine consignments across three regions with differing outreach structures, and practice was absorbed into routine two-hour refresher sessions.

Monitoring was kept deliberately simple and tied to routine systems, with four dials specified and thresholded to compare performance across the three regions and identify where adaptation might be needed. (This was noted as fidelity at or above 80% correct completion, on-time doses gaining at least five percentage points relative to matched districts, equity gaps narrowing for marginalized families, and cost per additional fully vaccinated child remaining below ten dollars). A single decision gate was recorded in plain language: expand no further unless at least two of the three regions meet all thresholds after two quarters – otherwise, pause, fix, and retest.

Vertical Scaling Plan — Pathway selection

In parallel, the team selected the vertical routes on the plan: policy integration into EPI, service-level integration through SOPs, supervision forms, and ordering, along with institutionalization via a ministry circular, an EPI budget line, and a training module. Legislative change was noted as not required for immediate adoption.

Vertical Scaling Plan — Categories of change

Each category of change refers to a specific part of the system that must adjust for the intervention to become routine—such as policy, financing, logistics, training, supervision, or information systems. Each was assigned an owner and a corresponding document or process to update.

D E P TH SCALE & SHARE

DEPTHS TOOLKIT

Vertical Scaling Plan

Intervention: Appointment reminder card

Type of Scale-Up

[X] Policy integration into an existing programme – Embed the intervention into an established programme, service, or policy framework.

[X] Service-level integration – Incorporate the intervention into routine service delivery platforms (e.g., adding to health worker protocols, school curricula).

[X] Institutionalisation – Make the intervention part of standard operating procedures, guidelines, or job descriptions, with dedicated budget lines.

Category of change	Change needed	Describe specific changes needed
Policy	Yes	Postcards must be included in MoH vaccination operational guidelines and demand-generation policies.
Political commitment	Yes	Ministers and senior officials must endorse and champion postcards as a national priority.
Legal change	No	Not required.
Regulations, norms, guidelines	Yes	Add postcards to immunisation SOPs, supervision checklists, and reporting templates.
Financing and budgets	Yes	Secure a dedicated budget line for postcard design, printing, and distribution.
Logistics	Yes	Integrate postcard supply with existing vaccine distribution chains.
Management information systems	To be defined	May need small adjustments to record postcard distribution and track fidelity.
Supervision	Yes	Supervisors must include postcard use in standard monitoring visits.
Staff evaluation, performance incentives	To be defined	Postcard delivery may be linked to outreach performance, but this needs assessment.
Training curricula and approaches	Yes	Add postcards to outreach worker training modules and refresher sessions.
Health workforce changes	No	No new staff needed; uses existing staff.
IEC materials	Yes	Postcards must be co-branded with MoH/UNICEF logos, translated into multiple languages, and pictorially adapted for low-literacy groups.
Other	Yes	Mechanism for periodic redesign (to prevent message fatigue) should be established.

Dissemination

- **Decision-makers:** MoH, UNICEF, WHO, district directors.
- **Approach:** Policy briefs and cost summaries for MoH/donors; peer learning for managers; job aids for workers; evidence briefs/webinars for multipliers.
- **Message:** Low-cost, equity-enhancing, easy to integrate.

Costs & resources

- **Costs:** At scale, $0.20 per postcard; 500,000 postcards annually (≈$100,000.)
- **Efficiency:** Economies of scale achievable with bulk printing and bundling with vaccine shipments.
- **Resources needed:** Donor co-financing in early phases; eventual transition to MoH budget line. Communications experts for design adaptation.

Monitoring & evaluation

- **Monitoring:** Supervisors check postcard use; immunisation data track uptake/timeliness; spot surveys assess caregiver recall.
- **Key indicators:** % postcards distributed correctly; % increase in on-time vaccination; equity gap reduction; cost per additional child vaccinated.
- **Evaluation:** Quasi-experiment to assess impact on vaccination uptake

Policy and guidelines required a ministry circular and revisions to EPI operational guidance to formalize the postcard use in routine immunization. **Financing** called for creating a dedicated EPI budget line, initially supported by donor funds and later absorbed into domestic financing. **Logistics** required an ordering code for reminder cards, bundling with vaccine shipments, and quarterly district stock counts. **Training** added a short module on postcard use to in-service refreshers and pre-service materials. **Supervision** integrated a tick-box for correct postcard completion on monthly visit forms, and recognition for high-performing teams. **Information systems** remained light, using existing immunization indicators to track progress, while a simple schedule for periodic postcard redesigns was added to sustain user interest over time.

Vertical Scaling Plan — Decision-making and dissemination

Dissemination for vertical decisions followed the plan precisely. EPI leadership received a two page brief and a short decision deck, while a low-key site visit preceded tabling at the National Immunization Technical Advisory Group. Nothing was staged: a caregiver showed the postcard at home and a nurse read out the next date already written. Requests to the Ministry were narrow and concrete — issue the circular, activate the budget line, add the supervision tick-box, and switch on the ordering code. NITAG endorsed, and the circular was signed the following week.

Learn more

This field guide introduces practical tools to help teams assess readiness, build support, and plan for scale. Scaling, however, is a complex field in its own right, drawing on insights from implementation science, systems change, political economy, and organizational management. No single guide can capture all of the perspectives, lessons, or strategies available. For that reason, this section points to additional resources for those who want to go deeper — whether to explore the frameworks introduced here in further detail, or to broaden their understanding of scaling challenges and approaches.

"I want to understand more about John List's *Voltage Effect.*"

"Voltage Effect" details potential 'voltage drops' that occur when interventions that look strong in a pilot lose effectiveness or cost-efficiency as they scale (see Appendix 1 for more detail).

"I want a quick, structured check of scalability before investing."

The Intervention Scalability Assessment Tool (ISAT) by Milat et al. gives policymakers and implementers a practical checklist to judge readiness for scale across domains like evidence strength, costs, fidelity, and context.

"I want to complement the Scale phase of the DEPTHS process with another scaling alternative."

WHO/ExpandNet: Nine Steps for Developing a Scaling-Up Strategy is a field-tested roadmap covering what to scale, who will adopt it, resource needs, partnerships, and monitoring. Pair it with ExpandNet: Beginning with the End in Mind to design pilots that keep future scale (budgets, roles, and data systems) in view from day one.

Another option is the BehaviourWorks Scale-up Toolkit, which helps teams to map core vs. adaptable elements, plan adoption pathways, anticipate voltage drops, and choose scale tactics (e.g., staged rollout, new channels). It's a good operational bridge between Horizontal and Vertical Scaling Plans.

"I want robust implementation frameworks to understand the context before implementation."

Teams can use the CFIR (Consolidated Framework for Implementation Research), which helps to identify barriers and enablers across the intervention, along with the setting, people, and processes involved. EPIS (Exploration, Preparation, Implementation, Sustainment) adds a phased view and "bridging factors" between systems and services. For busy teams, use CFIR to structure your risks/assumptions log, and EPIS to plan when and how adaptations will happen.

"I need to design adaptations without losing the core."

The Dynamic Adaptation Process (DAP) shows how to plan and document adaptations while protecting core functions. It's useful when your Horizontal Adaptation Plan calls for changes across languages, channels, or staffing models.

"I want to plan for long-term fit, not just launch."

The Dynamic Sustainability Framework (DSF) argues that context changes, and so must interventions. Use it to set up light, continuous improvement loops (e.g., quarterly fidelity + outcome reviews) during scale so the intervention stays effective as conditions shift.

"I need to monitor reach, fidelity, and maintenance during scale."

RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) gives five plain-English outcomes you can track with routine data. Combine it with PRISM if you need extra prompts on context (workflows, leadership, data systems) that shape those outcomes. Use these to choose a small set of indicators for your scale dashboards.

"I want to track and communicate broader public value."

The Translational Science Benefits Model helps to document community, clinical, economic, and policy benefits beyond the main effect size. It's useful for donor updates, budget justifications, and policy briefs during vertical scale-up.

"I need to turn evidence into decisions and products."

FHI 360's Research Utilization Framework maps a practical route from evidence to action—stakeholder mapping, product design, and institutionalization. For writing and planning, the PRB Research Translation Toolkit offers templates for policy briefs, stakeholder plans, and research-to-action roadmaps. The Value-Added Research Dissemination Framework gives a simple, end-to-end view of how to package and share findings so they are used.

UNICEF also offers a strong base of internal expertise and resources to support scaling. Staff with direct experience, along with internal documents and guidance, are available to help teams navigate this stage. Drawing on these assets can make scaling more achievable, especially considering that it is often the most complex part of applying behavioural science in practice.

RESOURCE	WHAT IT OFFERS & HOW IT'S USEFUL FOR SCALING
Social & Behaviour Change Programme Guidance: All Tools (UNICEF SBC Guidance)	A set of internal guidance, framing documents, and tools for SBC. Helps with designing, implementing, and iterating behaviour change programmes — useful when scaling to ensure consistent approach and quality.
Implementation Research Resources (UNICEF)	Practical guidance and case studies on how to use implementation research (IR) to adapt and scale interventions in real-world settings. Useful for identifying barriers, testing adaptations, and informing sustainable scale-up.
Scaling Up Child Protection: A Framework (Vol. 1, Vol. 2)	Provides a conceptual and practical roadmap for scaling programmes, policies, and services in child protection. Steps include building consensus, assessing scalability, scaling strategy, implementation, monitoring, and adaptation.
Implementation Research Compendium (UNICEF)	A collection of case studies from nine countries showing how implementation research (IR) helps in real-world settings, especially for adaptation and scale.
Scaling innovation for every child (UNICEF)	The UNICEF Innovation Group's principles and guidance for scaling innovations, particularly digital, within UNICEF's ecosystem.
Tools and ethics for applied behavioural insights: Basic toolkit (UNICEF Knowledge Summit)	Ethics and tools in applying behavioural insights. Useful for ensuring scaling is done responsibly, and that as reach increases, any risk of harm, unintended outcomes, or mis-application is managed.

Appendices

Appendix 1: Five Voltage Drops according to John List[27]

VITAL SIGN	WHAT IT MEANS	EXAMPLE
False positives	The pilot appeared effective, but the result was not real or replicable in the first place.	With DARE (Drug Abuse Resistance Education), early school pilots reported encouraging shifts in attitudes and intentions, with the programme spreading to thousands of schools. Later independent studies that tracked actual behaviour found little to no reduction in drug use, and in some cases, small backfire effects. The early "success" came from short term, self reported measures that did not hold up when larger groups and longer follow-ups were tested. The lesson is to confirm results with real outcomes and replication before wide rollout.
Population representativeness	The pilot worked for a sampled population, which is vastly different from the population at scale.	A preschool curriculum in a Chicago suburb raised test scores more for Hispanic families than for white or Black families, partly because Hispanic households in the setting were more often multigenerational, with grandparents able to help when parents were unavailable. If the team had scaled indiscriminately, it would have over-invested for groups that benefited less – as such, knowing who the intervention worked for would have guided targeting and adaptation.
Spillovers	If an intervention affects groups other than those sampled, then the impact at scale will not be like the impact in the initial test.	Positive spillovers had appeared in a preschool case, where children who were not enrolled still improved by playing with classmates who were – scaling would have amplified impact.
Supply side	Even if benefits persist at scale, 'diseconomies of scale' can cause a voltage drop if expansion of the programme causes costs to rise disproportionately.	A team planned for costs and delivery capacity as the intervention developed, deliberately designing the Chicago Heights curriculum so that ordinary communities relying on "average" teachers could deliver the intervention. This made the model more financially and operationally scalable.
Context representativeness	The intervention worked in a particular situation that was too different from the world at scale.	The early childhood programme succeeded with a limited number of excellent teachers, but hiring 30,000 such teachers would not have been feasible. Similarly, Early Head Start struggled when parents with spare time, their key ingredient, were no longer available. In short, the situation that enabled the pilot's success would not have scaled.

27 List, J.A. Optimally generate policy-based evidence before scaling. Nature 626, 491–499 (2024). https://doi.org/10.1038/s41586-023-06972-y

Appendix 2: Operational dimensions that most often break at scale.

Pilots often work because they benefit from extra funding, supervision, or ideal conditions that no longer exist when scaled up. The table below lists the operational areas that most often break under routine conditions — from staffing and training to supplies, data, and community access. Reviewing these areas helps to spot where a pilot's advantages can't be assumed, and where adjustments will be needed to make scale possible.

OPERATIONAL DIMENSION	CONSIDERATIONS
People and workload	• How many staff are on duty per session • Who accomplishes which tasks (nurses, clerks, outreach workers) • Usual sick leave or staff leaving, and how gaps are covered • Protected time versus competing duties • Time required for each task • Shift patterns and overtime rules • Reliability of volunteers • Any limits on what each role is permitted to do • How quickly replacements can be hired
Training and onboarding	• Length and style of training (classroom, or on the job) • Practice time and what counts as a pass • Number of trainers per group • Languages used, and whether materials are easy to read • Who trains whom (central team or local supervisors) • How often refreshers occur • Job aids provided • How new or rotating staff are brought up to speed • Whether training fits into regular staff meetings and budgets

Table continues on following page.

OPERATIONAL DIMENSION	CONSIDERATIONS
Supervision and QA	• How often supervisors visit, and for how long • How many staff each supervisor supports • Tools used by staff (simple checklists, observation forms) and how feedback is given • How problems are raised, by whom, and how quickly they are fixed • How many sessions are observed • Whether mentoring or peer support is available • Whether supervision actually occurs, given travel time and workload
Service delivery pattern	• Days and hours clinics are open • How often outreach occurs, and how it is planned • How long a typical session runs, and whether it starts on time • Client flow (busy times, queues, fast track for very young children) • Appointment and no-show rates • Which services are offered together (e.g., child growth checks with vaccines) • Clashes with other events, like campaigns or market days • Seasonal changes in demand
Caseload and coverage	• Average number of clients per day or session • Share of new vs. return visits • Size and spread of the catchment area • Busy seasons • Drop-off between doses • Number of individuals requiring follow-up each week • Coverage targets and recent trends • Gaps between neighbourhoods or groups • How often target lists are updated, and how accurate they are

Table continues on following page.

OPERATIONAL DIMENSION	CONSIDERATIONS
Supplies and logistics	• Who forecasts needs, and how • How often deliveries arrive, and typical waiting times • Storage space in rooms and in fridges or freezers • How temperatures are tracked • Agreed buffer stock levels • How often items run out, and for how long • Wastage and whether the soonest-expiring stock is used first • Whether items come in complete sets (syringes, safety boxes, cotton, gloves) • Storage security and tidiness • How accurate stock records are
Transport and access	• Travel time to outreach sites • Whether vehicles are available, and how they are booked • Fuel rules and budgets • Breakdowns and maintenance • Road and weather issues • Reliability of public transport • Checkpoints or security stops • Back-up options, like motorbikes or hired cars • Cost and approvals needed for trips

Table continues on following page.

OPERATIONAL DIMENSION	CONSIDERATIONS
Data and IT	• Main source of data (paper register, simple tally, electronic register) and how it advances through the system • How often reports are sent, and whether they arrive on time and complete • How many people share a device • Network coverage and any offline options • Charging and battery issues • Who helps when devices or software fail, and how quickly • Common errors (name mix-ups, duplicate records) • How permission and privacy rules are followed • Who accesses dashboards and how they are used in meetings
Payments and incentives	• Rules and rates for daily allowances, meals, or travel • Airtime or data support for phone calls or messaging • Any performance bonuses, what they are based on, and how quickly they are paid • Volunteer stipends or other recognition • Whether staff feel payments are fair • The length of paperwork processing, from claim to payment • How secure the funding source is
Communication channels	• Official methods of sharing decisions (memos, circulars, regular meetings) and how long it takes for front-line staff to hear about changes • Informal channels (WhatsApp groups) and who moderates them • Community channels (local radio, town announcements, faith groups) • Whether messages are easy to read and written in the correct languagesHow feedback from front-line staff reaches managers • How rumours are spotted and addressed

Table continues on following page.

OPERATIONAL DIMENSION	CONSIDERATIONS
Procurement and finance	• How long it takes to buy printing, protective equipment, stationery, or fridge parts • Whether there are standing agreements with suppliers or one-off purchases • Who signs off at each amount, and how long it takes • Limits on small cash • Spending bans at the end of a quarter or year • Supplier reliability and late payments • Exchange-rate or inflation risks • Tax or duty exemptions • Paperwork needed for financial checks
Policy and governance	• What approvals are needed (ministry notice, technical group sign-off, ethics if required) • Whether plans match current guidelines and written procedures • Any rules that limit data sharing or consent • Who decides at national, district, and facility levels • How accountability works • How often inspections or checks occur • Alignment with partner rules (UN/NGO)
Environment and security (if relevant)	• Curfews or movement limits • Patterns of conflict or violence • Disaster seasons (floods, storms, extreme heat) and back-up plans • Safe routes for outreach • Insurance or risk cover • Community acceptance and possible backlash • Staff stress and support • Contingency stock and mobile cold-storage options

Table continues on following page.

OPERATIONAL DIMENSION	CONSIDERATIONS
Equity or access constraints	• Languages and literacy levels • Disability access (ramps, signs, communication aids) • Gender norms that affect attendance or decision-making • Time and travel costs for caregivers • Need for identity documents and who lacks them • Status of refugees or migrants, and their entitlements • Cultural or religious calendars • Risks of stigma or discrimination • Tailored solutions, such as interpreters, flexible hours, or privacy during visits

Resources

1. Al-Ubaydli, Omar, and John A. List. 2022. "Scaling for results in development: Option C thinking." Issues in Science and Technology 38 (4). https://issues.org/scaling-option-c-al-ubaydli-list/
2. Al-Ubaydli, Omar, and John A. List. 2023. "Scaling for impact: The road ahead." Issues in Science and Technology. https://issues.org/scaling-option-c-al-ubaydli-list/
3. BehaviourWorks Australia. n.d. "Chapter 12: Making the most of an effective intervention." https://www.behaviourworksaustralia.org/the-method-book/chapter-12-making-the-most-of-an-effective-intervention
4. Centre for Economic and Social Studies (CESS), V. Reddy, and N. Jayakumar. 2011. Financing the WASH sector in India: Cost of provision and budget allocations (WASHCost Working Paper No. 12; CESS Working Paper No. 100).
5. Crider, Yoshika, Miki Tsuchiya, Magnifique Mukundwa, Isha Ray, and Amy J. Pickering. 2023. "Adoption of point-of-use chlorination for household drinking water treatment: A systematic review." Environmental Health Perspectives 131. https://doi.org/10.1289/EHP10839
6. Damschroder, Laura J., David C. Aron, Rosalind E. Keith, Susan R. Kirsh, Jeffery A. Alexander, and Julie C. Lowery. 2009. "Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science." Implementation Science 4 (1): 50. https://doi.org/10.1186/1748-5908-4-50
7. Karunamoorthi, K. 2014. "The counterfeit anti-malarial is a crime against humanity: A systematic review of the scientific evidence." Malaria Journal 13: 209. https://doi.org/10.1186/1475-2875-13-209
8. List, John A. 2024. "Optimally generate policy-based evidence before scaling." Nature 626: 491–499. https://doi.org/10.1038/s41586-023-06972-y
9. List, John A., F. Momeni, A. Peysakhovich, and D. Suskind. 2023. "A behavioral framework for scaling." Nature 624: 85–90. https://www.nature.com/articles/s41586-023-06972-y
10. List, John A., F. Momeni, and Yves Zenou. 2023. "The voltage effect in scaling." Nature 623: 963–970. https://doi.org/10.1038/s41586-023-06972-y
11. Luke, Douglas A., Ana A. Baumann, Benjamin J. Carothers, John Landsverk, Enola K. Proctor, and Aurelio J. Figueredo. 2018. "Forging a link between mentoring and capacity building in implementation science." Translational Behavioral Medicine 8 (4): 652–659. https://doi.org/10.1093/tbm/ibx071
12. Milat, Andrew J., Kirsten Lee, Katy Conte, Alethea Grunseit, and Luke Wolfenden. 2020. Intervention Scalability Assessment Tool (ISAT): A decision support tool for health policy makers and implementers. Sydney: Centre for Epidemiology and Evidence, NSW Ministry of Health. https://www.health.nsw.gov.au/research/isat
13. Ogutu, E. A., A. S. Ellis, K. A. Hester, K. Rodriguez, Z. Sakas, C. Jaishwal, C. Yang, S. Dixit, A. S. Bose, M. Sarr, W. Kilembe, R. Bednarczyk, and M. C. Freeman. 2024. "Success in vaccination programming through community health workers: A qualitative analysis of interviews and focus group discussions from Nepal, Senegal and Zambia." BMJ Open 14 (4): e079358. https://doi.

org/10.1136/bmjopen-2023-079358

14. Patient-Centered Outcomes Research Institute (PCORI). n.d. "Dissemination and implementation framework and toolkit." https://www.pcori.org/implementation-evidence/putting-evidence-work/dissemination-and-implementation-framework-and-toolkit

15. Sanders, Michael, Veerle Snijders, Michael Hallsworth, and Eamonn Kirkman. 2021. "Scaling: A behavioural science perspective on scaling up behavioural insights interventions." OSF Preprints. https://osf.io/preprints/osf/scd3k_v1

16. United Nations Development Programme (UNDP). 2013. Guidance note: Scaling up development programmes. New York: UNDP. https://www.undp.org/publications/guidance-note-scaling-development-programmes

17. University of Washington. n.d. "Implementation science resource hub." Seattle: University of Washington. https://impsciuw.org/

18. World Health Organization. 2010. Nine steps for developing a scaling-up strategy. Geneva: World Health Organization. https://iris.who.int/bitstream/handle/10665/44432/9789241500319_eng.pdf?sequence=1

19. World Health Organization. 2011. Beginning with the end in mind: Planning pilot projects and other programmatic research for successful scaling up. Geneva: World Health Organization. https://iris.who.int/bitstream/handle/10665/44432/9789241500319_eng.pdf

Conclusion

From insights to impact

Behavioural science builds on insights from psychology, behavioural economics, systems thinking, sociology, and human-centred design to explain the social, environmental, structural, and psychological factors that shape human behaviour. For international development and child rights, these insights challenge long-held assumptions about what drives change and reveal new, practical ways of supporting behaviours that are healthy, protective, and fair.

Applied behavioural science turns understanding into action. It helps to uncover what truly drives behaviour and grounds programmes in evidence rather than assumptions. By testing and learning systematically, teams can reduce waste, improve outcomes, and scale what works.

What behavioural science teaches us

Behavioural science offers several key lessons for UNICEF and its partners in international development and humanitarian action:

- **People are social.** Behaviour is shaped by relationships, norms, and shared identities — not just by information or intent. Programmes that build on social networks and community values are more likely to succeed.
- **Context matters.** Choices are influenced by systems and environments. Sustainable change happens when those environments make healthy and fair options the easiest ones.
- **Effective action depends on evidence, not intuition.** Perceptions and judgments are prone to error. Assumptions about what works to change behaviour are often misguided. Testing ideas separates what works from what merely feels right. Measurement and experimentation reveal what drives change, for whom, and under what conditions.
- **Ethics and humility are essential.** Even well-intentioned efforts can do harm if not designed and tested with a duty of care. Ethical practice requires meaningful consent, respecting people's agency, protecting privacy, and ensuring that results and benefits are shared with communities.

A mindset for action

While applied behavioural science offers valuable tools and methods to generate stronger evidence, its true value lies in its mindset — a way of thinking about how change happens. It provides a structured lens for seeing, understanding, and solving problems, grounded in how people actually think, decide, and act within the real conditions of their lives and environments. By replacing assumptions with a systematic process for testing and learning, this mindset enables teams to design, refine, and scale solutions that work in practice, not just in theory.

The DEPTHS process translates this mindset into a practical approach for designing and improving behaviourally informed programmes. It offers a structured framework to move from insight to implementation, ensuring solutions are evidence-based, context-appropriate, and sustainable. To review:

- D **Define:** Clarify the problem, target behaviours, and scope to guide the work.
- E **Explore & Diagnose:** Gather and analyse evidence to understand the drivers, barriers, and contextual factors shaping target behaviours.
- P **Prototype Designs:** Co-create and refine solutions with users to ensure they are practical, relevant, and evidence-based.
- TH **Test Hypotheses:** Measure whether interventions drive the desired behavioural change and generate evidence to guide adaptation or scale.
- S **Scale:** Build support and plan to expand proven solutions for sustained impact.

Embedding a culture of learning

DEPTHS is more than a sequence of steps — it is a shift in how we work. It builds experimentation and reflection into every stage of programme design, making evidence generation and use a normal part of delivery. This helps teams test assumptions early, learn quickly, and adapt continuously. In short, DEPTHS turns uncertainty into insight and ideas into impact.

The process is not meant to be applied rigidly or mastered all at once. Many teams already use similar methods and principles as part of good practices in intervention design and programming. What the DEPTHS process adds are structured moments and tools to deepen curiosity, clarify problems and mechanisms, and continuously test and improve.

Start small: ask sharper questions, test one assumption, or co-create new solutions with a few community members. These small steps build habits of curiosity, experimentation, and learning that grow stronger over time.

The BIRD Lab supports UNICEF and partners with training, technical support, and mentorship—from diagnosing behavioural challenges to prototyping, testing, and scaling solutions. By sharing experiences and results, UNICEF teams and implementing partners can build collective capacity in applied behavioural science and deliver greater impact for children.

Looking ahead

Behavioural science continues to evolve, deepening our understanding of how people make decisions across cultures and contexts. Yet much of today's evidence still comes from WEIRD societies, while experiences from the Global South remain underrepresented. By building evidence where it is scarce, adapting methods to real-world settings, and sharing lessons widely, UNICEF can help shape a more inclusive and globally relevant field.

This is not a final destination but an ongoing journey — one that advances with every new insight and every community we work with. By staying curious, evidence-driven, and ethical, we can design programmes that are not only effective but transformative — creating lasting change for every child.

Key terms

This list collects all the key terms used throughout the field guide.

Core terms used in the Introduction Chapter

- **Applied behavioural science (ABS):** Leveraging evidence from different fields studying human behaviour to solve real-world problems by intentionally changing people's decision making for the better. It offers practitioners three things: an evidence base, a process to understand and change behaviours, and robust evaluation and testing methods.
- **Availability bias:** A mental shortcut that leads people to overestimate the likelihood of events that are easy to recall, such as plane crashes or shark attacks.
- **Behavioural economics:** A field that examines how people actually make decisions, often in ways that differ from rational choice models.
- **Behavioural intervention:** Practical application of evidence-based insights from behavioural science to address specific programme and policy challenges
- **Behavioural science:** The empirical study of how people think, decide, and act, with an emphasis on observation, systematic experimentation, and evaluation.
- **Biases and heuristics:** Mental shortcuts that the mind uses to make rapid decisions.
- **Choice architecture:** The practice of designing how choices are presented to people to influence their decisions.
- **Cognitive bias:** A systematic pattern in how people interpret information or make decisions, shaped by mental shortcuts, emotions, and social or contextual influences.
- **Confirmation bias:** The tendency to notice, value, and believe information that supports one's existing beliefs.
- **Context:** The internal and external setting that affects people's decision-making and behaviour. Contexts are made up of the interaction between people (demographics, norms, personal history) with a place (physical environment, available resources) at a specific moment (including the person's mindset and the local stimuli) and with a specific history (culture).
- **Default bias:** When people are presented with options, the default choice carries tremendous influence.
- **DEPTHS Define phase:** The first step of the DEPTHS process, focused on clearly identifying the specific problem and behaviours to address.
- **DEPTHS Explore phase:** The second step of the DEPTHS process, using research to understand the sociocultural, environmental, and psychological factors influencing behaviour.
- **DEPTHS Prototype phase:** The third step of the DEPTHS process, which involves co-creating and pre-testing potential solutions with the people who will use them.
- **DEPTHS Scale phase:** The final step of the DEPTHS process, which involves expanding effective interventions to new contexts (horizontal scale) and embedding them into policies and systems (vertical scale).
- **DEPTHS Test Hypotheses phase:** The fourth step of the DEPTHS process involves piloting interventions in real-world conditions to generate evidence and adapt designs.
- **DEPTHS:** An acronym for a systematic process developed by UNICEF to apply behavioural science. It stands for Define, Explore and diagnose, Prototype and design, Test Hypotheses, and Scale.
- **Dual process theory:** A model describing that our minds use two systems simultaneously for decision-making: a fast, automatic one (System 1) and a thoughtful, deliberate one (System 2).
- **Empathy gaps:** The difficulty people have in predicting how they or others will think, feel, or behave in a different emotional or situational state.

- **Evidence base:** The collection of research findings, data, and other forms of information used to support a particular recommendation, decision, or practice.
- **Fundamental attribution error:** The tendency to overemphasize personal traits (like character) and underemphasize situational factors when explaining others' behaviour.
- **Habits:** Behaviours that are automatically triggered by environmental cues and are performed with little conscious thought.
- **Human-centered design (HCD):** An approach that grounds the design process in the lived experiences of end users to ensure solutions are relevant and context-specific.
- **Intention-action gap:** The common phenomenon where people genuinely intend to do something but fail to follow through with action.
- **Mental limits:** People have limited time and energy, and most of their mental effort goes into a small number of complex decisions each day.
- **Mental shortcuts:** They are heuristics that people rely on to make decisions quickly, such as looking at what others are doing, following what feels good, or defaulting to familiar routines. They are useful because they help us navigate complex environments efficiently, but they can also lead to systematic biases or errors in judgment.
- **Present bias:** The tendency to accord more weight to immediate costs and benefits than to those that occur in the future.
- **Robust evaluation methods** (e.g. RCTs): RCTs and other experimental methods to measure behavioural impact.
- **Social and behavioural change (SBC)**: Is a strategic, evidence-based process that uses communication to promote and sustain positive social norms and behaviour change in individuals and communities. It goes beyond simply raising awareness and aims to address the complex mix of individual, social, and structural factors that influence people's choices and actions. The ultimate goal is to foster an environment that supports long-term, positive change for the well-being of communities.
- **Status quo bias:** Is the tendency to stick with existing conditions even when there are strong reasons to choose otherwise.
- **System 1:** The fast, automatic, impulsive, and reactive thinking system. It is likened to an "elephant" that is powerful but susceptible to environmental cues and biases.
- **System 2:** The slow, reflective, attentive, and deliberate thinking system. It is likened to a "rider" who can steer the elephant but requires conscious effort.
- **Systems thinking:** An approach that looks at the structural and systemic factors influencing behaviour to find pathways to scale solutions.
- **WEIRD societies:** An acronym for Western, Educated, Industrialized, Rich, and Democratic societies. The term highlights that people from these societies, who represent only 12% of the world's population, make up as much as 80% of study participants in psychology research.

Core terms used in Chapter 1: Define

- **Behaviour Tree:** A simple visual map that links a desired outcome to the key audiences and the specific behaviours that influence that outcome.
- **Behaviour:** A specific, observable action that a person takes, as opposed to their thoughts, feelings, or beliefs.
- **Drivers:** Social, psychological, and environmental factors—such as beliefs, norms, infrastructure, or rules—that influence whether a behaviour occurs.
- **Impact-Feasibility Prioritization Matrix**: A tool used to narrow focus by comparing behaviours against two criteria: impact (how much changing the behaviour would contribute to the desired outcome) and feasibility (how realistic it is to influence or change the behaviour).
- **Leverage Point Analysis:** A process that uses a system map to identify leverage points, which are aspects of a system that, if changed, would have a high impact on the desired outcomes.

- **Leverage points:** Specific parts of a system where a strategic behaviour change could lead to a significant improvement in outcomes.
- **Negative feedback loop** (Balancing loop): A process where a system responds to change by counteracting or reversing it, helping the system maintain stability or return to its original state.
- **Positive feedback loop** (Reinforcing loop): A loop that amplifies change by creating a self-reinforcing cycle that accelerates the original shift and pushes the system further in the same direction.
- **Primary audiences:** Individuals or groups who experience a problem firsthand, who directly influence the targeted outcome through their actions, or who will most likely benefit from a solution.
- **Problem Definition and Outcome Statement**: A tool used to clearly specify the problem, the outcome(s) of interest, who to engage, and which population groups to target.
- **Project Canvas:** A one-page summary document that outlines the intended outcome, leverage points, potential actors, and actions that will drive change, along with roles and responsibilities.
- **Secondary audiences:** Individuals or groups who indirectly shape the target outcome through their decisions, habits, or roles. They influence the environment, decisions, or access of the primary audience and can include gatekeepers, influencers, and decision-makers.
- **SMART goal:** A goal or outcome that is Specific, Measurable, Achievable, Relevant, and Time-bound.
- **Stakeholder and Audience Map:** A tool used to clarify which individuals are most interested in or concerned by a problem, along with their influence.
- **System Map:** A graphic representation showing the relationships between different factors in a system, including both drivers and barriers to success.
- **Systems thinking:** An approach that identifies different parts of a problem, their interactions, and how they influence each other over time.

Core Terms Used in Chapter 2: Explore & Diagnose

- **Barriers:** The specific psychological, social, structural, or environmental factors that block progress, prevent action, or make a desired behaviour more difficult to perform.
- **Behavioural Drivers Model** (BDM): A conceptual guide developed by UNICEF that illustrates the multiple drivers of behaviour, reminding users that a full understanding requires attention to psychological, social, and environmental influences.
- **Behavioural Mapping and Diagnoses:** A tool that traces the small, observable steps (micro-behaviours) along a behavioural pathway and identifies the barriers and enablers behind them.
- **Behavioural Profiles:** A tool used to synthesize insights about the psychological, socio-cultural, and contextual drivers behind a specific behaviour for a target population, illuminating who the key actors are, what shapes their choices, and what challenges they navigate.
- **Beneficence** (with regards to the Belmont Report): An ethical principle from the Belmont Report about actively seeking to maximize potential benefits while minimizing potential harm.
- **Capability** (in COM-B): The psychological and physical capacity to engage in an activity, including knowledge, skills, and mental faculties.
- **Cluster sampling:** A method where the population is divided into groups (clusters), some clusters are randomly chosen, and everyone within the chosen clusters is included in the study.
- **COM-B framework:** A model that breaks behaviour down into three essential components: Capability, Opportunity, and Motivation, used to diagnose the factors that may be enabling or inhibiting a behaviour.
- **Convenience sampling:** A method that selects people who are easiest to reach and willing to participate.

- **Desk research:** The process of reviewing existing data, literature, and insights to build a foundational understanding of the context, behaviours, and populations. It is also referred to as a "literature review."
- **Disproportionate stratified sampling**: A method that divides a population into strata but intentionally selects more people from certain smaller or more important groups.
- **Drop-off points**: Moments in a behavioural journey where a person stops progressing toward the target behaviour, either by delaying, skipping, or never completing a step.
- **Enablers**: Factors that are already helping a person move toward a desired behaviour or could be reinforced to make the behaviour easier, more likely, or more appealing.
- **Feasibility–Impact Matrix**: A tool used to assess and prioritize behavioural barriers and enablers based on how impactful addressing them would be and how feasible it is to do so.
- **Field Observations**: The act of watching how behaviours and interactions unfold in their natural settings.
- **Focus Group Discussions**: Facilitated discussions with a small group to uncover norms, perceptions, and shared experiences.
- **Grey literature**: Reports, briefs, or working papers that are not published in traditional academic journals but may contain valuable programmatic or partner insights.
- **In-depth semi-structured interviews:** One-on-one conversations guided by a set of questions to explore a set of experiences and motivations from a particular point of view.
- **Information power:** A concept used in qualitative sampling that focuses on the value of the information each participant brings to a study, rather than relying on arbitrary rules of saturation to determine sample size.
- **Institutional Review Board (IRB):** An ethics body that provides approval for research to ensure it is conducted responsibly and with appropriate safeguards.
- **Judgemental purposive sampling:** A method where the researcher uses their expertise to select people who are especially knowledgeable or relevant to the topic.
- **Justice** (with regard to the Belmont Report): An ethical principle from the Belmont Report about promoting fairness and ensuring that the risks and benefits of research are equitably distributed.
- **Meta-analyses:** A statistical method, also considered a "very strong" type of evidence, that combines the results of multiple scientific studies to derive a single, more precise estimate of effect.
- **Micro-behaviours:** The smaller, observable steps into which a complex behaviour is broken down for analysis.
- **Mixed-methods approach:** A research strategy that draws on the complementary strengths of both qualitative and quantitative techniques to construct a richer, more holistic understanding.
- **Motivation** (in COM-B): The reflective processes (like beliefs and intentions) and automatic processes (like emotional responses and habits) that drive behaviour.
- **Mystery User:** A trained researcher who simulates a real user experience to assess a process and identify hidden barriers.
- **Opportunity** (in COM-B): The external conditions that make the behaviour possible, including environmental, socio-cultural, and structural factors.
- **Primary research questions:** Clear and behaviourally-informed research questions aligned with gaps identified during the desk review, focusing on uncovering why behaviours occur or fail to occur.
- **Primary research:** The process of organizing and implementing fieldwork (e.g., recruitment, training, data collection) to fill evidence gaps identified after desk research.
- **Proportionate stratified sampling:** A method that divides a population into subgroups (strata) and selects participants from each subgroup in the same proportion as they appear in the whole population.
- **Qualitative methods:** Research techniques (such as interviews, focus groups, and observations) that help uncover the "why" behind behaviours by exploring barriers and drivers.

- **Quantitative methods:** Research techniques (such as surveys) that help measure the "what" by identifying patterns, frequencies, and relationships across a larger population.
- **Quota purposive sampling:** A method that ensures a certain number of people from specific categories are included in the sample.
- **RCTs** (Randomized Controlled Trials): A type of experimental study considered a "strong" or "moderate" form of evidence where participants are randomly assigned to different groups to test an intervention.
- **Research protocol:** A detailed, step-by-step document describing exactly how complex, resource-intensive, or ethically sensitive research will be carried out, often required for ethical review.
- **Respect for persons** (with regard to the Belmont Report): An ethical principle from the Belmont Report acknowledging the autonomy of individuals and providing extra protection to those with diminished capacity.
- **Sampling:** The process of selecting a subset of individuals from a broader population to participate in research.
- **Saturation:** The point in a literature review where further searching is unlikely to yield new insights, signalling that the most critical information has likely been attained.
- **Service or administrative records:** Routinely collected data from platforms like health systems, education records, or registration logs.
- **Simple random sampling:** A method where every person in a group has an equal chance of being selected.
- **Snowball sampling:** A method that starts with a few participants who then refer others, useful for finding hard-to-reach populations.
- **Social desirability bias:** The tendency for participants to offer responses they believe are expected, appropriate, or socially acceptable.
- **Social network mapping:** A method to identify influencers, social expectations, and approval/disapproval dynamics around behaviours.
- **Surveys:** Structured questionnaires used to gather quantitative data from a larger population.
- **Systematic reviews:** A type of evidence considered "very strong" that involves a comprehensive review and synthesis of multiple high-quality studies on a specific topic.
- **Systematic sampling:** A method that selects every nth person from a list after starting at a random point.

Core Terms Used in Chapter 3: Prototype & Design

- **Behavioural Insights Team** (BIT): Also known as the "Nudge Unit," it is the organization that developed the EAST framework to apply behavioural science to public policy.
- **Co-creation:** A participatory design process that involves community members, frontline workers, and other key stakeholders in generating and shaping solutions.
- **Cognitive misers:** A concept from cognitive psychology stating that humans naturally prefer to use simpler and less mentally taxing ways of thinking and making decisions.
- **COM-B model:** A framework for understanding behaviour that posits that for any behaviour to occur, a person must have the Capability, Opportunity, and Motivation.
- **Decision-makers and long-term implementers:** The individuals or entities, like program managers or government officials, who approve, influence, or sustain an intervention.
- **Deliverers:** The people responsible for implementing the intervention on the ground, such as health workers or teachers.

- **Design ethics:** The ethical considerations focused on ensuring an intervention itself is fair, inclusive, and sensitive to unintended consequences for the people it affects.
- **Design Provocations:** A creative technique that uses "what if" questions to challenge, deepen, and strengthen existing intervention ideas.
- **EAST framework:** A behavioural science framework based on four principles—Easy, Attractive, Social, and Timely—used to design effective interventions.
- **End users:** The individuals or groups who will directly experience or use the intervention.
- **Fidelity (of a prototype):** The level of detail and functionality of a prototype, or how closely it resembles the final version of the intervention.
- **How might we questions:** Structured prompts used to translate research findings into actionable opportunities, sparking creative and practical intervention ideas.
- **Ideation:** The creative process of brainstorming and generating a wide range of behaviourally-informed ideas for interventions.
- **Intermediate outcomes:** The short-term cognitive or emotional shifts (e.g., increased intention, better planning) that occur after an intervention is delivered and precede the final behaviour change.
- **KAP** (knowledge, attitudes, practices) surveys: Surveys used to measure a population's knowledge, attitudes, and practices on a specific topic.
- **Micro-behaviour:** A specific, small action that a person is expected to adopt, which contributes to a larger, overall behaviour change goal.
- **Outputs:** The immediate, observable, and quantifiable results of an intervention's activities, which show that the intervention is being delivered as intended (e.g., number of reminder cards distributed).
- **Parallel testing (A/B testing)**: A user testing method where multiple versions of a prototype are presented to users to compare reactions and preferences.
- **Primary behavioural outcome:** The specific, measurable action or behaviour that the intervention ultimately aims to change.
- **Prototype:** A low-cost, preliminary, and tangible version of an intervention (e.g., a sketch, model, or role-play) created to gather feedback from users and stakeholders.
- **Prototyping:** The process of creating a preliminary, simplified version of an intervention to be tested with users for early feedback.
- **Prototype types:** Common formats used to make ideas tangible—storyboards, role-play, paper mock-ups, physical models, experience prototypes, concept videos, service blueprints, SMS/message scripts.
- **Research ethics:** The formal ethical principles that govern how people are studied, including informed consent, safeguarding, privacy protection, and IRB approval.
- **Theory of Change (ToC):** A comprehensive explanation that maps out how and why an intervention is expected to cause the desired behavioural outcomes and overall impact.
- **User-testing methods:** Lightweight ways to gather feedback—informal walkthroughs, role-plays/simulations, small-group discussions, observation-only testing, implementer dry runs, and lightweight pilots.

Core Terms Used in Chapter 4: Test Hypotheses

- **Attrition**: Loss of participants over time; can reduce power and bias estimates, especially if differential by arm.
- **Back Checks:** A quality control method where a subset of survey respondents are revisited to re-ask a few key questions to verify the accuracy and honesty of the original data collection.
- **Backdoor path:** In a Directed Acyclic Graph (DAG), an alternative, non-causal pathway between an intervention and an outcome that can introduce bias. Causal inference aims to "close" these paths to isolate the true effect.
- **Baseline measurement:** Data collected before the start of an intervention to enable comparison with follow-up outcomes.
- **Before-and-after comparison:** A simple but weak evaluation method that measures outcomes before a program and again after it, attributing any change to the intervention.
- **Causal effect:** The difference between the potential outcome with the treatment and the potential outcome without it (Y1−Y0).
- **Causal pathways:** The processes through which an intervention leads to observed outcomes, measured to understand why and how effects occur.
- **Causation:** The relationship where one event or action directly causes another to happen. Determining causation is the core challenge in evaluation.
- **Comparison group:** A group of individuals that is as similar as possible to the intervention group but does not receive the intervention. It is used to estimate the counterfactual.
- **Confounding factors:** External influences, other than the intervention being studied, that can affect the outcomes and potentially distort the results, making it difficult to isolate the true impact of the intervention.
- **Contamination:** Control participants access the intervention or its effects, weakening contrasts between arms.
- **Correlation:** A relationship where two things happen or change together, but one does not necessarily cause the other.
- **Cost-Benefit Analysis (CBA):** A systematic approach to estimate the strengths and weaknesses of an intervention by comparing all its costs to its benefits, which are typically expressed in monetary terms.
- **Counterfactual:** The hypothetical scenario of what would have happened to the participants in an intervention if they had not received the intervention. It serves as the benchmark for measuring true impact.
- **Directed Acyclic Graphs** (DAGs): Visual diagrams that represent the causal relationships between different variables. They are used to identify potential confounding factors and design more rigorous evaluations.
- **Effect size:** The magnitude or size of the change that an intervention is expected to produce.
- **Endline measurement:** Data collected after treatment to estimate impact and track persistence of effects.
- **Endogeneity:** A situation in statistics where the relationship between an intervention and its outcome is distorted because other unobserved or unaccounted-for variables are influencing both.
- **Experiments** (Randomized Controlled Trials / RCTs): A type of impact evaluation where participants are randomly assigned to treatment and control groups to measure the causal effect of an intervention.
- **False negatives / Type II error:** Failing to detect a true effect due to insufficient sample size or weak study power.
- **Fade-out effects** (Decay): When behaviour or outcome improvements weaken or disappear over time.
- **Fundamental problem of causal inference:** The dilemma that it's impossible to observe both the actual outcome (with the intervention) and the counterfactual outcome (without the intervention) for the same individual at the same time.
- **High-Frequency Checks** (HFCs): The regular and timely review of incoming data during collection to quickly identify and address errors, inconsistencies, or potential fraud.

- **Impact Evaluation:** An evaluation approach that aims to determine the causal effect of a program or intervention, answering the question, "Did the program cause the observed change?"
- **Institutional Review Board** (IRB): An ethics committee that reviews research methods involving human subjects to ensure that they are ethical and that participants' rights and welfare are protected.
- **Instrumental Variables:** A quasi-experimental method that uses a variable (the "instrument") that influences participation in a program but is not directly related to the outcome. This helps to isolate the program's causal effect in the presence of confounding factors.
- **Implementation fidelity:** The degree to which program delivery follows the intended design, important for interpreting evaluation results.
- **Intention-to-Treat (ITT) effect:** An analysis method in randomized trials where participants are analyzed in the groups to which they were originally assigned, regardless of whether they actually received or completed the intervention.
- **Matching Methods:** A quasi-experimental technique that creates a comparison group by matching individuals who received an intervention with similar individuals who did not, based on observable characteristics.
- **Multiple testing adjustments:** Statistical procedures (e.g., family-wise error or FDR controls) to account for testing many outcomes/subgroups.
- **Non-Equivalent Comparison Group:** An evaluation design that compares an intervention group to a group that did not receive the intervention during the same time period but was not randomly assigned. It is an improvement over a before-after design but is still vulnerable to selection bias.
- **Potential outcomes framework:** A formal statistical structure for thinking about causal impact by considering two potential outcomes for each unit (person, household, etc.): the outcome if they receive the treatment (Y1) and the outcome if they do not (Y0).
- **Power calculations:** A statistical procedure used before a study begins to determine the minimum sample size needed to reliably detect a specific effect size, if it truly exists.
- **Pre-analysis plan (PAP):** A document created before data analysis begins that specifies the hypotheses to be tested and the exact statistical methods that will be used. It helps prevent bias from "data fishing" or selective reporting.
- **Process Evaluation:** An evaluation approach that focuses on how a program was implemented, examining its delivery, fidelity, reach, and the operational context. It answers the question, "How was the program delivered?"
- **Protocol registration** (trial registration): Publicly posting core protocol elements before data collection to timestamp pre-specification and reduce reporting bias.
- **Quasi-Experiments:** Evaluation methods used when randomization is not feasible. They use statistical techniques to create a comparison group to estimate an intervention's impact but carry a higher risk of bias.
- **Randomization:** The process of randomly assigning participants to either receive an intervention (treatment group) or not (control group). It is considered the "gold standard" for establishing causality because it creates statistically equivalent groups.
- **Rapid-cycle testing:** Iterative, small-scale tests that measure quickly, learn, and refine before larger evaluations.
- **Representativeness:** Ensuring that the sample reflects the target population sufficiently to generalize results.
- **Robustness (sensitivity) checks:** Pre-specified analyses that test whether results hold under alternative specifications/assumptions.
- **Selection confounders:** Pre-existing differences between the group that participates in a program and the group that does not, which can bias the results. This includes self-selection and targeting bias.
- **Social desirability bias:** When respondents modify answers to appear favourable or acceptable.

- **Statistical power:** The probability that a study will detect an effect when there is a true effect to be found. A common target for power is 80%.
- **Sub-group analysis:** Examining whether an intervention works differently for different groups (e.g., by gender, age, or region).
- **Third variable problem (or Common cause confounding):** A situation where a hidden, third factor influences two other variables at the same time, making them appear correlated without a direct causal link.
- **Time-based confounders:** Changes that occur over time, independent of the intervention, such as seasonal variations or pre-existing trends, which can be mistaken for the intervention's effect.
- **Triangulation:** Using multiple data sources or methods to validate findings and reduce measurement bias.
- **Unit of randomization:** The level at which allocation occurs (individual, household, community/cluster, facility), chosen to balance inference, spillovers, logistics, and power.

Core Terms Used in Chapter 5: Scale

- **Context representativeness:** A cause of "voltage drop" where the pilot situation was too different from the real-world conditions at scale.
- **CORRECT attributes:** Attributes that make interventions more likely to succeed at scale. They are: Credible, Observable, Relevant, Relative advantage, Easy to install and understand, Compatible, and Testable.
- **Decision gate.** A pre-agreed threshold that determines whether to progress, pause, or adapt during staged scale-up.
- **Environment Assessment:** The key activity of assessing environmental readiness for scaling.
- **False positives:** A cause of "voltage drop" where a pilot's result was not real or replicable.
- **Feasibility at scale:** Whether an intervention can be delivered reliably within routine systems, budgets, and staffing levels.
- **Horizontal scaling:** The expansion or replication of an intervention across new geographic areas or population groups.
- **Option C thinking:** A mindset that pushes teams to ask "will it continue to work under real-world conditions at scale?" rather than just "does the idea work?".
- **Population representativeness:** A cause of "voltage drop" where the pilot population is too different from the population at scale.
- **Scaling:** A deliberate effort to increase the impact of successfully tested innovations so that more people benefit and so that policies and programmes can evolve on a lasting basis.
- **Sharing:** The intentional dissemination of insights, evidence, and lessons so that others—whether communities, partners, or policymakers—can adopt, adapt, and multiply impact in their own contexts.
- **Spillovers:** A cause of "voltage drop" where an intervention affects groups other than those sampled, changing the net impact at scale.
- **Stress Test:** The key activity of pressure-testing interventions under realistic conditions.
- **Supply side voltage drop:** A cause of "voltage drop" where the costs of expanding a program rise disproportionately.
- **Unintended effects:** Harmful or unexpected outcomes that may emerge when interventions are expanded.
- **Vertical scaling:** The process of embedding an intervention within national or sub-national systems through policy, budgets, or service delivery structures to institutionalize it. Vertical scaling increases the number of (similar) people who receive the intervention.
- **Voltage Drop:** The decline in effect size when an intervention moves from pilot to large-scale implementation.
- **Voltage Effect:** The key activity of testing whether an intervention's effectiveness will hold at scale.

Project team (UNICEF):

Ukasha Ramli, Benjamin Hickler, Madeline Quinlan, Floriza Gennari

Internal Advisory Group (UNICEF):

Dorina Andreev-Jitaru, Cassia Ayres, Helena Ballester Bon, Noel Migabo Balume, Viviane Melo Bianco, Rudrajit Das, Michelle Dynes, Francine Elvia Ganter Restrepo, Amaya Gillespie, Karen Greiner, Gloria Lihemo, Ken Limwame, Taitos Matafeni, Mario Mosquera, Markel Roberto Mendez Hernandez, Neha Kapil, Deepa Risal Pokharel, Johary Randimbivololona, Sergiu Tomsa, Vincent Petit, Massimiliano Sani, and Siddartha Shrestha.

Acknowledgements:

This field guide is part of a package of applied behavioural science resources developed in collaboration with many colleagues across UNICEF. Special thanks are due to Social and Behaviour Change teams from the following Country Offices, who provided feedback on early versions of training resources and tools: Bolivia, Cambodia, Ethiopia, Ghana, Honduras, India, Indonesia, Lao PDR, Malaysia, Mozambique, Myanmar, Nicaragua, Nigeria, Philippines, Senegal, Sri Lanka, Tanzania, Timor-Leste, and Zambia.

UNICEF works in the world's toughest places to reach the most disadvantaged children and adolescents and to protect the rights of every child, everywhere. Across 190 countries and territories, we do whatever it takes to help children survive, thrive and fulfil their potential, from early childhood through adolescence. And we never give up.

UNICEF Innocenti – Global Office of Research and Foresight tackles the current and emerging questions of greatest importance for children. It drives change through research and foresight on a wide range of child-rights issues, sparking global discourse and actively engaging young people in its work. UNICEF Innocenti equips thought leaders and decision makers with the evidence they need to build a better, safer world for children. The office undertakes research on unresolved and emerging issues, using primary and secondary data that represent the voices of children and families themselves. It uses foresight to set the agenda for children, including horizon scanning, trends analysis and scenario development. The office produces a diverse and dynamic library of high-level reports, analyses and policy papers, and provides a platform for debate and advocacy on a wide range of child-rights issues.

UNICEF Behavioural Insights Research and Design Laboratory, DEPTHS:
An Applied Behavioural Science Field Guide, March 2026.

Technical Partner Consortium: This field guide was researched, written, tested and designed by a partnership of the best minds in behavioural science: Led by Common Thread, with Busara and Firsthand.

Disclaimer: This is a working document. It has been prepared to facilitate the exchange of knowledge and to gather feedback. The text has not been edited to official publication standards and UNICEF accepts no responsibility for errors. The statements in this publication are the views of the author(s) and do not necessarily reflect the policies or the views of UNICEF. The designations in this publication do not imply an opinion on legal status of any country or territory, or of its authorities, or the delimitation of frontiers.

© United Nations Children's Fund (UNICEF), March 2026

Except for the UNICEF logo, photos and where another © appears, this work is licensed under CC BY 4.0. The UNICEF logo and enclosed photos may only be used in exact copies of the publication. Where another © appears, you must ask the listed copyright holder to use the item outside exact copies of the publication.

To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

BIRDLAB

BEHAVIORAL INSIGHTS RESEARCH AND DESIGN