

'Improving health and wellbeing: a guide to using behavioural science in policy and practice'

Evaluating Behaviour Change Interventions



Authors:

Dr Alice Cline, Senior Behavioural Science Specialist, Public Health Wales
Dr Nicky Knowles, Principle Behavioural Science Specialist, Public Health Wales
Jonathan West, Head of Behavioural Science, Public Health Wales
Lucia Homolova, Senior Public Health Research Officer, Public Health Wales
Dr Charlotte Grey, Public Health Evaluation Lead, Public Health Wales
Dr Esther Mugweni, Deputy Head of Evaluation, Public Health Wales
Ashley Gould, Programme Director, Behavioural Science Unit

Review Date: March 2025

We are really interested to hear about your experience of applying behavioural science – in the spirit of all share – all learn, so please do share your quick reflections on using this tool with us via : PHW.behaviourchange@wales.nhs.uk

Contact details:

Behavioural Science Unit, Policy and International Health, WHO Collaborating Centre on Investment for Health & Well-being
Public Health Wales.

Email: PHW.behaviourchange@wales.nhs.uk

Website: <https://phwwhocc.co.uk/bsu>

Twitter: @BSU_PHW

Behavioural Science Unit:

The Public Health Wales Behavioural Science Unit provides specialist expertise on behavioural science, and enables and supports the application of it, to improve health & wellbeing in Wales. The Unit is part of the Policy and International Health Directorate World Health Organisation (WHO) Collaborating Centre on Investment in Health and Wellbeing.

For further information, or support around the application of behavioural science to improve and protect health and wellbeing in Wales please get in touch.

Mae'r ddogfen hon ar gael yn Gymraeg / This document is available in Welsh

Some aspects of this tool have been previously published, and are owned by others. Their content has been translated, with retention of some of the originally published language and design.

The Central Evaluation Team has been established within the Research and Evaluation Division within the Knowledge Directorate at Public Health Wales. The team supports the design and delivery of robust evaluation across PHW and its wider strategic partners.

ISBN: 978-1-83766-363-7

© 2024 Public Health Wales NHS Trust.

Material contained in this document may be reproduced under the terms of the Open Government Licence (OGL) provided it is done so accurately and is not used in a misleading context.

www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

Acknowledgement to Public Health Wales NHS Trust to be stated. Many thanks to the individuals and teams who supported this work by sharing their insights and experiences.

Copyright in the typographical arrangement, design and layout belongs to Public Health Wales NHS Trust.

Purpose of this Tool

Applying behavioural science helps to optimise the impact of activities, services, policies, and communications where behaviours are on the critical path to success. Developing such interventions that account for, or aim to influence, behaviour involves several steps, as shown in Figure 1; these steps are fully described in the 'Improving health and wellbeing: a guide to using behavioural science in policy and practice' which you can find [here](#).

To support each step, we have released a series of deep dive tools. In [tool 1](#), we worked through how to develop a behavioural specification. In [tool 2](#), we focused on how to collect behavioural insights, something we call a behavioural diagnosis. In [tool 3](#), we outlined how to map our behavioural insights onto the COM-B model, in order to identify and prioritise intervention options. In [tool 4](#), we looked at different implementation options that could be used to deliver an intervention. In [tool 5](#), we worked through how to identify and apply Behaviour Change Techniques when developing interventions.

This final tool in the series supports step number 6, evaluating behaviour change interventions. Whether you're developing a new intervention or adapting interventions which are already being delivered, you will most likely have already been considering or planning how best to implement and evaluate.

Written in collaboration with the Central Evaluation Team at Public Health Wales, this tool provides an outline of the key points to consider when planning how to evaluate a behaviourally informed intervention. This tool is not intended to be a comprehensive guide to evaluation *per se*, many such guides and tools have been provided elsewhere, and so we've provided links to further reading throughout. By the end of this tool, we'll have developed a plan to help ensure our application of behavioural science to our intervention(s) has the best chance of success, is ethically sound and that we can understand the effectiveness, adding to the existing evidence-base.



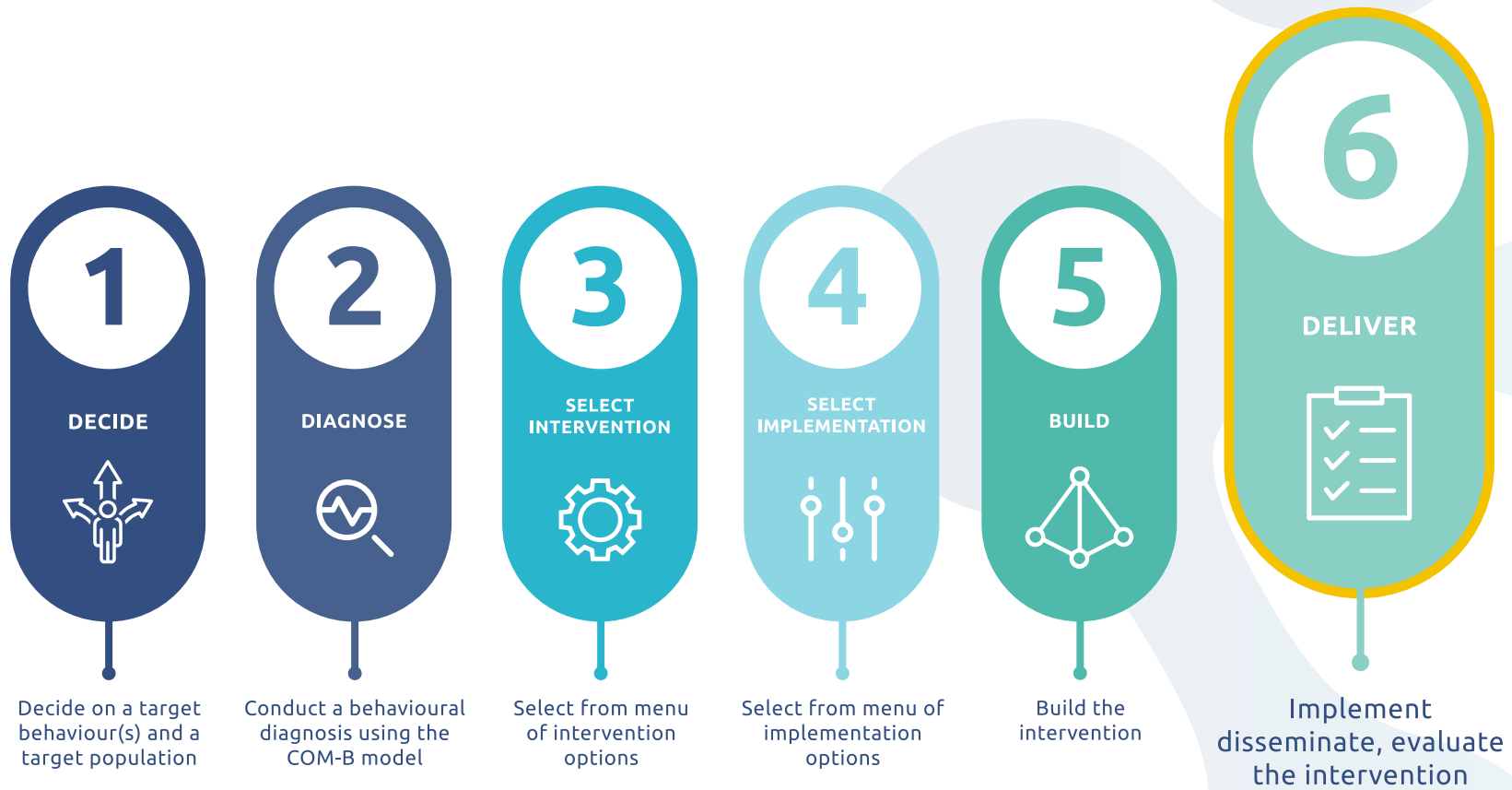


Figure 1. Six steps to developing an intervention

Principles of Applying Behavioural Science

In our overarching guide, we introduced the key principles to applying behavioural science in policy and practice. Steps five, six and seven are especially relevant when thinking about implementation, dissemination, and evaluation.

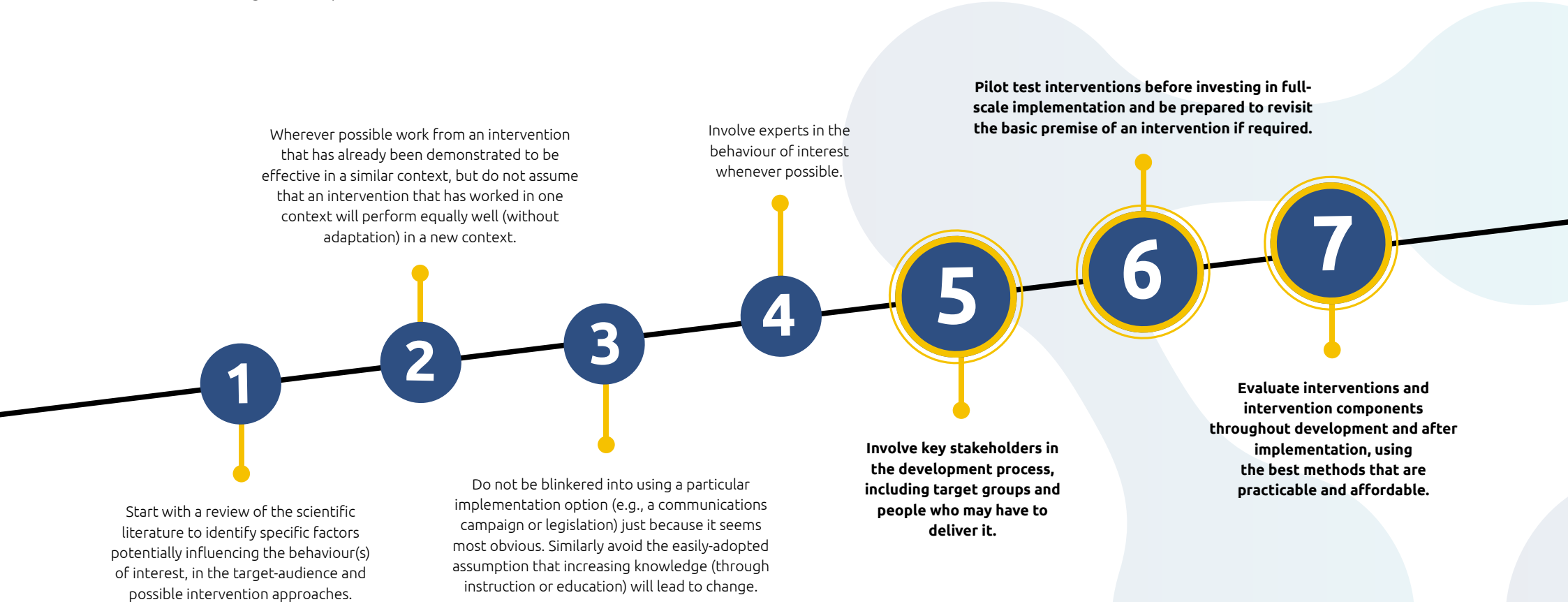


Figure 2. Principles to apply during the development process

Evaluating Behaviourally Informed Interventions

Good quality evidence on the impact of behaviour change interventions is critical to inform decision-making and build knowledge. Evaluation helps us to develop a deeper understanding of whether or not our intervention (or adaptation of/combination of interventions) has been effective in changing behaviour; for who (which population groups), how and why it has worked. It also helps us to understand the intended and unintended consequences of the interventions. Evaluation involves assessing the intervention against a number of evaluation criteria as defined in evaluation frameworks and guidance (e.g., APEASE, which you can read about on page 8).

This information gained from evaluation can be used to improve, scale up, or replicate successful behaviour change interventions, as well as to develop new ones. Sometimes, it may be necessary to stop or withdraw a behaviour change intervention as a result of the information gathered during an evaluation. Evaluation should ideally be planned in at the start of the intervention design stage, being clear about what learning the evaluation will provide and exactly what will be evaluated. Evaluation also has an accountability function, so the evaluation can help inform whether the intervention's objectives have been met and whether it is value for money. For example, in addition to health outcomes, we may want to understand how behaviours have changed, whether the determinants of the behaviours have changed or, whether the intervention was implemented as intended. You can read more about evaluations [here](#).

Evaluations can take a theory-based approach to explore the processes through which the intervention leads to outcomes. For example, gaining understanding of how and why change is brought about including through Behaviour Change Techniques, and the interaction with context that influence the effectiveness of the behaviour change intervention. Evaluations can also take an efficacy- or effectiveness-approach. For example, to understand the impact of the intervention on health and wellbeing outcomes, changes to the desired behaviour(s), and/or changes to the determinants of that behaviour. Behaviour change interventions can become increasingly complex in line with how many multiple interacting components (or interventions) there are, the range of behaviours and groups that are targeted, and how dynamic or adaptive the intervention is. Consideration of a range of approaches that can be used to evaluate complex interventions can be helpful, see [here](#) for further guidance.

The scale and complexity of your evaluation will depend on a range of factors ranging from what you want to learn from the evaluation to practical considerations such as the time and resources.

Behaviourally informed evaluations focus on changes to behaviour, and/or changes to the determinants of behaviour.

Different Types of Evaluations

Evaluations can vary in type and methods used, often depending on the questions we want answered and the stage of intervention and/or policy being evaluated. Evaluation at an early implementation stage (formative) can be used to further develop the intervention and fix implementation problems so that it is more likely to be successful. Evaluation at a later stage (summative) can help to determine whether the intervention has been a success or not. Ideally, interventions would be evaluated throughout implementation using different and complimentary types of evaluations, but in reality, the type(s) and stage(s) of evaluation conducted will depend on a variety of factors including resources available and how you intend to use the outputs. We've listed the different types of evaluation below:

Feasibility studies and evaluability assessments can be used to support your thinking at an early implementation stage; they typically involve literature reviews on similar topics, theory of change development, and stakeholder involvement. You can read more about feasibility studies [here](#) and you can read more about evaluability assessments [here](#). In some cases, it may be appropriate to adapt an existing intervention to a new context. The [ADAPT guidance](#) outlines questions to consider if you're thinking of doing so. Whether you're implementing something new, or adapting something that already exists, it's important to spend time at this stage by trialling phased implementation(s) and pilot testing where resources allow.

Outcome evaluations work to provide evidence on the short- and medium-term changes that occurred, whilst **impact evaluations** aim to collect evidence on the longer-term effects. Both types of evaluation can be theory-based (without a comparison group, e.g. realist evaluation or contribution analysis); experimental (e.g. Randomised Control Trial); natural/quasi experimental (with a non-randomised comparison group e.g. interrupted time-series analysis); and observational (e.g. pre-post, cross-sectional, cohort, and case-control studies); each with different strengths and limitations. There are guidance and frameworks to support your evaluation, such as [MRC Guidance for natural experiments](#). You can find out more about outcome evaluations [here](#).

Process evaluation involves using data from various sources including document reviews and primary data collection using qualitative and/or quantitative methods. Process evaluation can be conducted as a stand-alone study at an early or later stage of the intervention or alongside outcomes, impacts or economic evaluations. This type of evaluation can help you understand if your behaviour change intervention, and your behaviour change techniques, are being delivered as intended (e.g. what is implemented and how), how does your intervention produce change and how does context affect implementation and outcomes. There are evaluation criteria in guidance and frameworks already in place to help guide your process evaluation, such as the [RE-AIM Framework](#) and the [MRC Guidance for Complex Interventions](#). You can read more about process evaluations [here](#).

Economic evaluations can help to establish the costs and benefits of an intervention. They can provide particularly useful information to any policy or decision makers or health professionals about the value and use of resources and how they should be allocated. You can read more about them [here](#). In addition to measuring financial 'value for money', it is also important to understand and measure the social or public value of policies/intervention, including their wider social, economic and environmental outcomes and impacts. The social return on investment methods (SROI) is a participatory, often beneficiary-led approach to measure and quantify the value created by a programme or intervention in social, economic and environmental terms. SROI can capture what matters most to people, families, communities, funders and all relevant stakeholders, as well as to society, wider economy and planetary sustainability. This can help inform and strengthen the case for spending and investment in population health and prevention. More information and the unique Social Value Database for Public Health can be found [here](#).

Pilot Testing and Implementation

From a behavioural science perspective, spending time on pilot testing and phased implementation is particularly important before fully implementing a new or adapted behaviour change intervention. Behavioural interventions often involve shaping peoples' choices and habits, and sometimes it might be necessary to intervene in ways that restrict choices or are unpopular, so it is an important ethical principle to ensure that we have fully taken the target population's views and preferences into account. Using something like the APEASE criteria can help us to systematically consider implementation-imperatives, such as if the intervention is acceptable to the target population who will be receiving it, and the stakeholders who are delivering, funding, or involved in it.

The APEASE criteria:

Affordability	<ul style="list-style-type: none"> • Is the proposed intervention affordable? • Has the pilot/phase implementation met cost expectations? • Is the intervention likely to be affordable when scaled up? 	Acceptability	<ul style="list-style-type: none"> • Is/was the pilot intervention acceptable to the target population? • Do the stakeholders involved in delivering, designing, implementing and/or funding the intervention satisfied that it's acceptable? • Will the intervention be acceptable to the target population and other key stakeholders when delivered at scale?
Practicability	<ul style="list-style-type: none"> • Can the intervention be delivered as designed, through the means intended? • Will this be possible when the intervention is scaled up/ adapted for a larger population? 	Spill-over effects	<ul style="list-style-type: none"> • Are there likely to be any unwanted side-effects or unintended consequences when delivered at scale? • Did the pilot/phased implementation identify any unintended consequences (positive or negative)?
Effectiveness and cost-effectiveness	<ul style="list-style-type: none"> • Is the intervention likely to be effective when delivered within a real-world setting? • Has the pilot/phased implementation demonstrated effectiveness in achieving the intended outcome(s)? • Will this effectiveness likely remain when scaled up? • What is the ratio of effect, when compared to cost? 	Equity	<ul style="list-style-type: none"> • To what extent may the intervention increase or decrease existing inequalities? • Did the pilot/phased implementation show any evidence of increasing/decreasing inequalities?

**Adapted from the Behaviour Change Wheel, A Guide to Designing Interventions*

In addition to the target population, the behaviours of stakeholders involved in an intervention are often critical to its success. Building on the COM-B model (Mitchie et al 2013), it's important to understand the capabilities, opportunities, and motivations of those who will be involved in supporting, delivering, and promoting the intervention. By considering this in advance, we can address any gaps in their capability, opportunity, and motivation to ensure that all stakeholders are able to effectively perform their role in implementing the intervention. For example, do your stakeholders:

- Know their role and what they need to do?
- Have the required skills?
- Understand the purpose and importance?
- Have support from their peers/leaders/wider organisation?
- Feel confident about their role within the intervention?
- Have the required equipment/time/supportive policy/strategic environment?

We should be prepared to revisit the intervention during the development process should the testing indicate that the intervention isn't likely to have the desired effect. Learning from the pilot testing can then inform decision making around next steps which could be one of the following:

- Adopt: Proceed with full scale implementation
- Adapt: Refine the intervention based on learning from the pilot before proceeding
- Abandon: Do not proceed any further with the intervention

Often, it may be necessary to test, adapt and test again multiple times before the intervention is ready for full-scale implementation.

A common mistake is to undertake pilot testing but with a strong bias toward proceeding. Very often we hear of behaviour change interventions being described as 'promising' when an objective evaluation of the available data does not support such a view. In addition to the usual performance and operating constraints we work with, other biases that might lead to premature implementation include: sunk cost fallacy, the reluctance to abandon a strategy or course of action because it has been heavily invested in, even when abandonment would be more beneficial; confirmation bias, the tendency to search for or interpret information in a way that confirms one's preconceptions, and disregard information that does not support the initial opinion; self-serving bias, tendency for people to evaluate ambiguous information in a way beneficial to their interests.

When Should I Evaluate?

We can evaluate a behaviour change intervention at different stages. Evaluations can inform thinking before, during and/or after an intervention has been implemented. When we choose to evaluate will depend on the information that we would like to find out and what evaluation questions need to be answered.

Before/During Implementation

Why?

To inform intervention development and future evaluation design, in addition to identifying and reducing uncertainties in both.

For example, you may want to know:

- How will the intervention reach the target population?
- Is it practical to deliver the intervention to the target population, what about at scale?
- Is the intervention affordable, will it be affordable when/if scaled up?
- How will the intervention bring about change, can these assumptions be tested?
- Is the intervention acceptable to stakeholders?
- Is the intervention acceptable to the target population?
- Will there be side-effects as a result of the intervention?
- Is the intervention likely to increase/decrease existing inequalities, how?
- Is the intervention sustainable for the time required?
- What evaluation approaches should be considered at the next stage?
- What data needs to be collected to facilitate future evaluation?

Evaluation Type?

- Feasibility studies and evaluability assessments
- Formative process evaluation

How?

- Review the evidence available on previous evaluations of similar interventions
- Develop and refine your theory of change
- Pilot an intervention
- Test evaluation approaches in population of interest
- Establish what data is available or needs to be collected

When?

Suitable for new initiatives pilot programmes or trials, or programmes where you would like to understand who and how does it work, programmes that may want to be scaled up, and/or to understand how best the intervention can be adapted to new contexts.

Case Study: Toomey and colleagues (2017) wanted to establish the feasibility of delivery of a complex behaviour change intervention. The intervention was a self-management programme for people with osteoarthritis and chronic non-specific lower back pain, delivered in primary care physiotherapy. It consisted of 6-weekly sessions of 90-minutes and consisted of education and exercise. The aim of the evaluation was to (a) establish the fidelity of delivery of the complex behaviour change intervention – compared with usual physiotherapy (b) evaluate the fidelity of providing the SOLAS intervention and (c) explore the factors that may have influenced these results. To answer the evaluation aims, a feasibility study was conducted. The team used mixed methods design to collate the necessary data, including: audio recordings, direct observations, self-reported checklist, semi-structured interviews with the physiotherapists. This allowed them to conclude that the intervention had been delivered as intended, with high fidelity. You can read more about the evaluation [here](#).

With the intervention that you'd like to evaluate, and your evaluation objectives in mind, use the space below to consider how you may go about answering the questions bullet pointed on page 10:



EDITABLE
TEXT BOX

During Implementation:

Why?

To understand the intervention implementation in practice and emerging outcomes.

For example, you may want to know:

- Is the intervention acceptable to the target audience that it's being delivered to?
- Is the intervention acceptable to the stakeholders who are delivering it?
- To what extent are stakeholders adopting the intervention?
- Is the intervention being delivered as intended?
- To what extent the behaviour change techniques being delivered?
- Are the individual behaviour change techniques being implemented as intended, or have they been modified in some way?
- Is the intervention reaching the target population as intended?
- Is there evidence that behaviour(s) are changing? Or, that the behavioural determinants are changing?
- Are there any unintended (positive or negative) consequences? What are they?
- Does the intervention need to be modified to achieve the desired behaviour change?

Evaluation Type?

- Process evaluation
- Outcome evaluation
- Economic evaluation
- Impact evaluation

How?

- Collect data whilst the intervention is taking place

When?

Suitable for learning about the intervention implementation and emerging outcomes. This information can inform whether the intervention should be modified to achieve the intended behaviour change.

Case Study: The All-Wales Diabetes Prevention Programme (AWDPP) had been designed to identify and support adults who are at risk of developing type 2 diabetes, identified through a Hb1Ac measurement between 42-47 mmol/mol. The programme offers a one-to-one 30-minute person-centred conversation focused on behaviour changes to physical activity and diet. These conversations are guided by the COM-B model, exploring a patient's capability, opportunity, and motivation to change their behaviour(s). In 2023, Swansea, Aberystwyth, and Bangor University (SABU) Consortium was commissioned to evaluate the implementation of wave 1 of the AWDPP. The key objectives of the evaluation were to: (a) conceptualise the intervention context, (b) understand the intervention delivery and implementation, (c) examine the mechanisms of impact and (d) explore the value of the intervention, in line with Wales Prudent Healthcare Principles. To answer the evaluation aims, formative process evaluation was undertaken. The team used mixed methods design to collect the necessary data, through surveys, interviews, focus groups, clinic observations, analysis of routine clinical and administrative data and reviews of programme documents. This allowed the team to gain insights into how the intervention was being implemented, whether or not implementation differed between health boards, how the intervention was being received by service-users. Overall the team gained a greater understanding of what worked well and what recommendations could be made in order to inform future roll outs. You can read more about the evaluation [here](#).

With the intervention that you'd like to evaluate, and your evaluation objectives in mind, use the space below to consider how you may go about answering the questions bullet pointed on page 12:

After Implementation:

Why?

To identify the outcomes of an intervention, helping to demonstrate impact and provide learning for future efforts, and collect retrospective data regarding design and implementation.

For example, you may want to know:

- Did the intervention change the behavioural determinants?
- How effective was the intervention at changing the target behaviour(s)?
- What changed and for who? e.g., who did the intervention work for/who not?
- Were the changes sustained over time?
- Can the intervention be attributed to the changes in behaviours?
- Were there any other intended/unintended changes to behaviours/outcomes/impacts?
- Did the intervention increase/decrease inequalities?
- Was the intervention cost-effective?
- What have we learnt from the intervention that we share with others?

Evaluation Type?

- Process evaluation
- Outcome evaluation
- Impact evaluation
- Economic evaluation

How?

- Collect follow-up data to compare with the data measured at baseline
- Gather data on the outcomes of the intervention

When?

Suitable for producing evidence on the outcomes obtained, and collecting retrospective data regarding the design and implementation of an intervention.

Case Study: The ‘Safetxt’ intervention was designed to reduce sexually transmitted infections and increase safer sex by (a) encouraging participants to correctly follow STI treatment instructions, (b) promote condom use with a new or casual partners and (c) encourage participants to obtain testing for STIs prior to unprotected sex. The content was delivered by 42-79 text messages over one year, tailored according to the type of infection, gender and sexuality. ‘Safetxt’ comprised educational, enabling, and incentivising behaviour change strategies. The intervention involved 12 behaviour change techniques in total, such as information about the consequences of a behaviour and anticipated regret. To evaluate the intervention, Free and colleagues (2023) aimed to establish and quantify the effect of the ‘Safetxt’ intervention on the incidences of chlamydia and gonorrhoea infection at 12 months, compared with a control group receiving usual care. A Randomised Controlled Trial was implemented. The primary outcomes included data on incidences of infection at 12 months, in addition to self-reported data on condom use, number of people telling their partners about infection and the number of people who went for STI testing prior to sex with new partners. The evaluation concluded that the educational messages did not reduce infections over 12-months. The messages slightly increased condom use and the number of people telling their partner about infection, but did not change the number of people testing for infection prior to having sex with a new partner. You can read more about the evaluation [here](#).

With the intervention that you’d like to evaluate, and your evaluation objectives in mind, use the space below to consider how you may go about answering the questions bullet pointed on page 14:

Conducting an Evaluation

To take a systematic approach to designing your evaluation, you can consider working through the steps outlined below. As always, the extent to which you complete each step will depend on the resources you have available to you. This is how we can begin to build on the information provided above, and apply it in practice.

STEP**1**

Obtain clarity on the behaviour you're trying to change

Before starting your evaluation, it's important to be clear on your target behaviour the behaviour that your intervention is targeting and your target audience. This is called a behavioural specification, if you haven't developed one yet, we advise taking a look at tool 1 in this series, which you can find [here](#).

Behavioural Specification:

Who:

What:

How:

When:

STEP
2**Create a theory of change (ToC)**

The next step is to create a theory of change, if you've completed tool 5 you'll already have developed your theory of change, if not we've provided an outline for you below.

Theories of change help to visually illustrate how an intervention will work, how it is expected to bring about change and for who. They can provide an overview of how and why behaviour change is expected to occur and include information about any underlying assumptions (i.e. external conditions that must exist in order to achieve the desired outcomes), risk, and context surrounding the intervention that impact on how outcomes are achieved. Theory of change is a theory-driven approach, often developed in collaboration with key stakeholders to gain insights and help refine it through consultations or workshops.

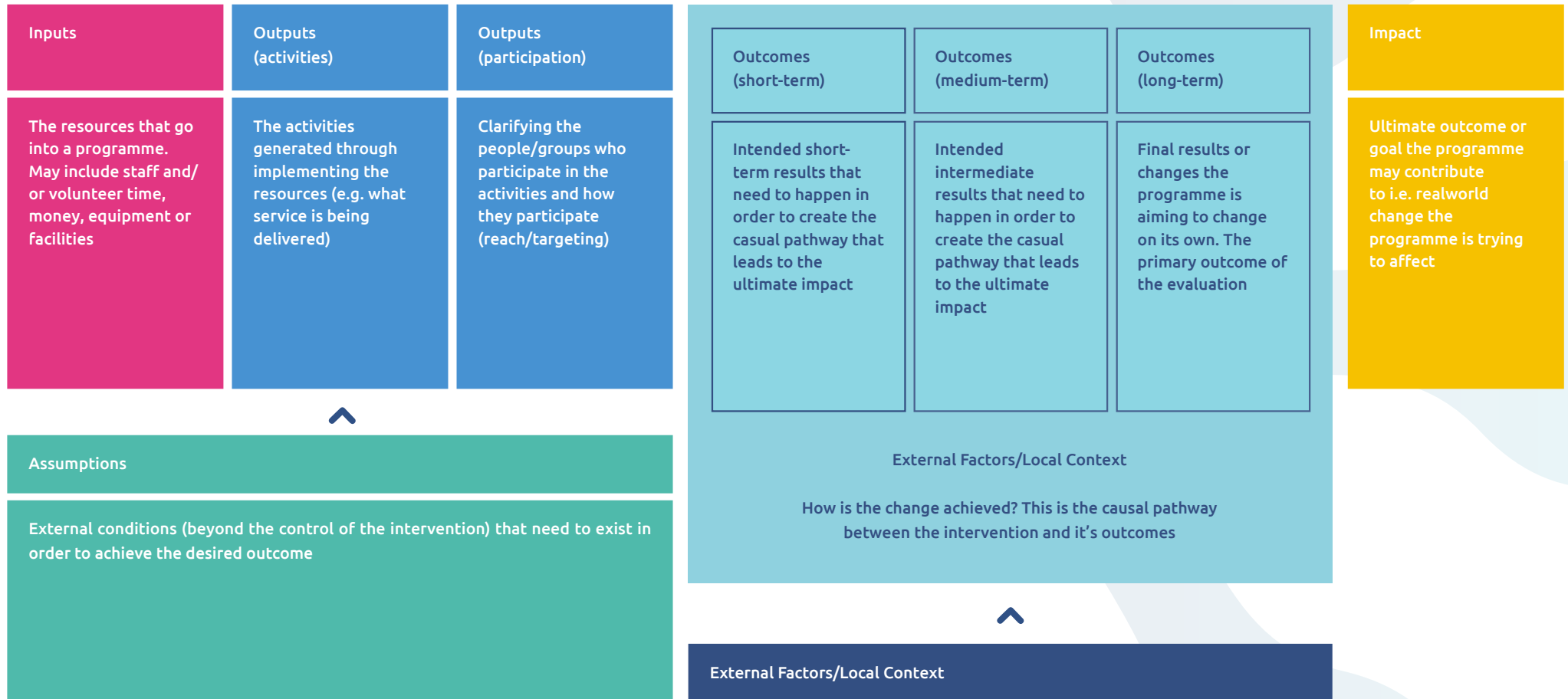
ToC can be very useful to help you guide your evaluation, as it can test the theory that has been developed and further contribute to refining it. ToC can be a particularly useful tool to help identify priority areas for the evaluation, inform the development of meaningful evaluation questions and identify key indicators for monitoring changes in your outcomes or what data you may require to test the theory.

There is no pre-set way to present a theory of change, but it is often visualised in a graphic format, which can vary from simple to more complex flow diagrams. The content of ToC will always be specific to the intervention. We have included below an example of ToC diagram:

<https://www.eif.org.uk/resource/developing-a-good-theory-of-change>

<https://evaluationhub.eif.org.uk/theory-of-change/>

*De Silva, M.J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C. and Patel, V., 2014. Theory of change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials*, 15(1), pp.1-13.*



When developing a theory of change, you usually start at the end and populate the diagram from right to left. This means starting with the behaviour change you'd like to see, followed by short to medium term outcomes that may be required to achieve the change. Finally, you then include details about the intervention activities required to achieve the desired outcomes.

To help inform your theory of change, you may want to consider the questions below:

Impact

- What is your target behaviour? It's important to think about your behavioural specification here.
- What will achieving your target behaviour lead to in terms of public health outcome(s)?

Outcomes (short, medium and long-term)

- Consider the behavioural determinants (barriers/facilitators) to your behaviour here, using tool 5 you can also think about Mechanisms of Action too.
- What behaviours may need to happen, in order for your target behaviour to happen? What order may they need to happen in?
- If the target behaviour is dependent on other behaviours, what may prevent those behaviours from happening?
- It's important to try to identify any other assumptions that are being made here (e.g., people will receive the invite letter through the post, open it and read it).
- Which Behavioural Change Techniques will be visible and present during the intervention?
- At this point, it's helpful to identify indicators that you can measure to understand if the short-, medium-, or long-term outcome has been achieved, how will you measure if behaviour has changed?

Outputs (participation)

- Who needs to participate, in order for the behaviour change to occur?
- Think about your behavioural specification here, who is your target population?
- How will they interact with the intervention? More specifically, how will they interact with the Behaviour Change Techniques that you've previously?

Outputs (activities)

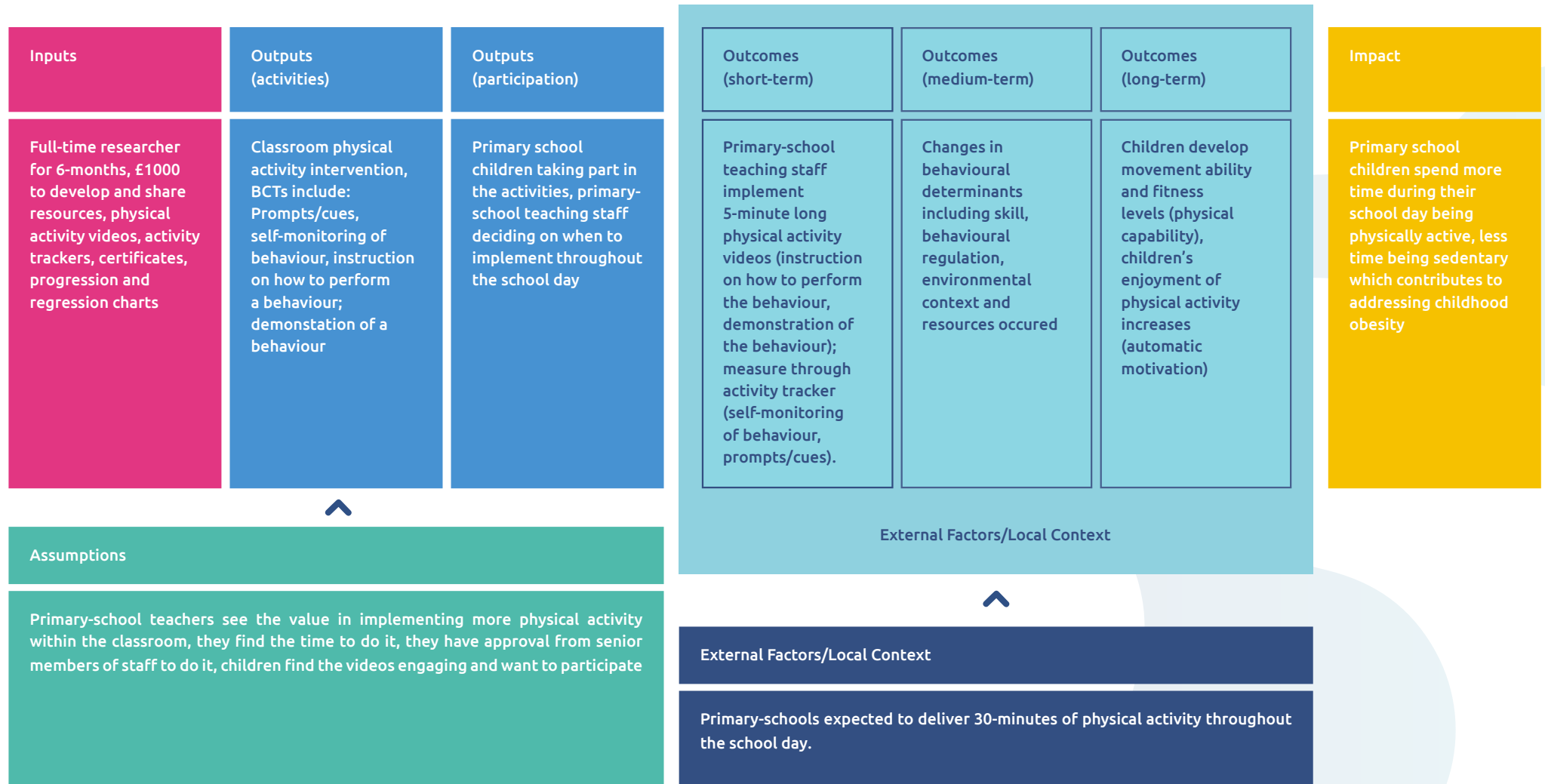
- What will the intervention do? You can consider the intervention functions that you've previously identified in tool 4 here.
- What's being delivered? You can consider what implementation types you identified in tool 5 here.
- What will the Behaviour Change Techniques look like as activities?
- Who will deliver it/how will it be delivered?

Inputs

- What's required in order for the intervention to take place?
- What resources are required (staff, time, money, equipment, facilities etc)?

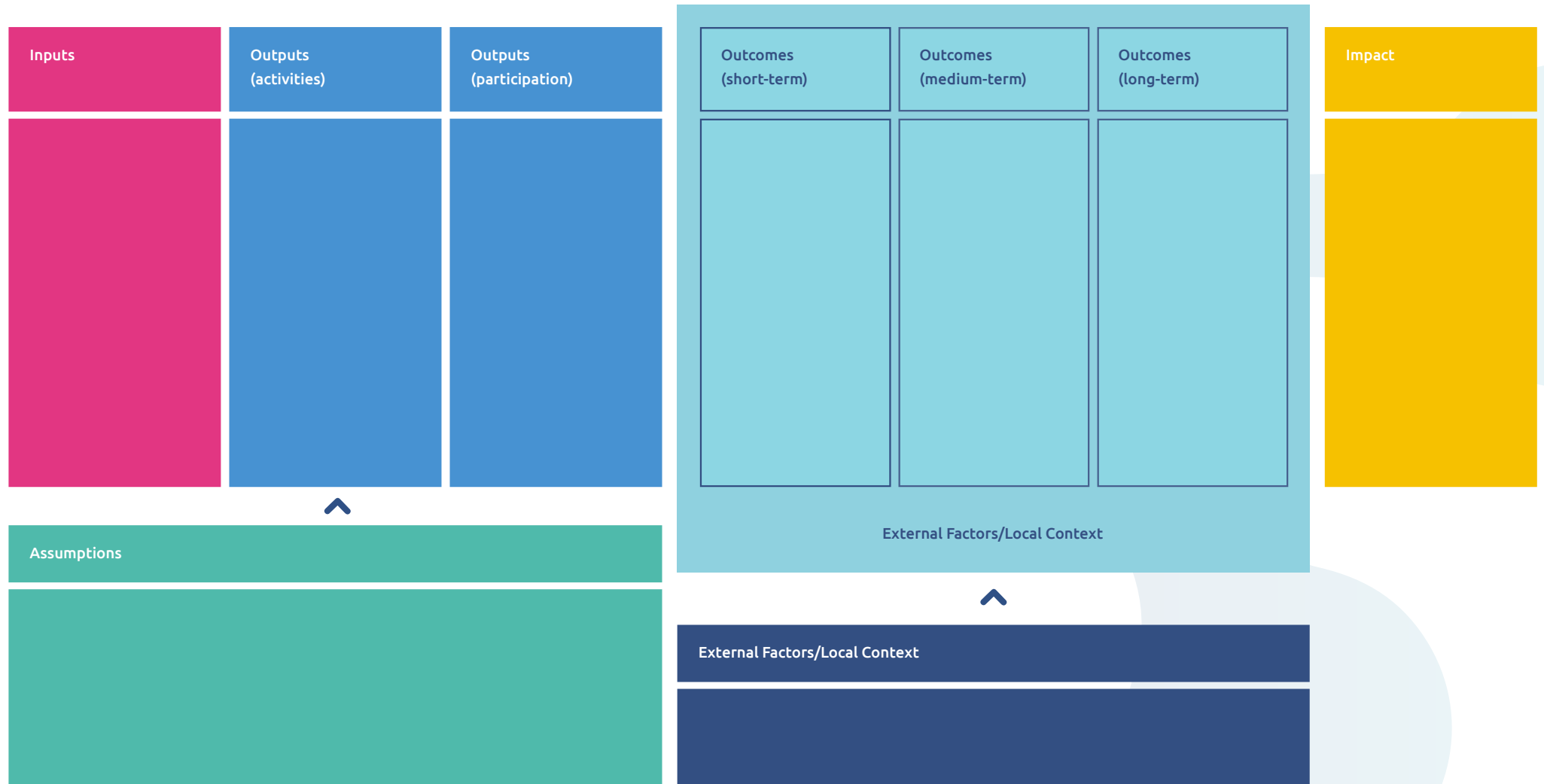
An Applied Example:

Here's an example of what a theory of change looks like in practice. This theory of change describes an intervention that was implemented within primary school classrooms, with the intent to increase children's physical activity levels.



Template:

We've provided a blank theory of change template for you to use, of course your theory of change doesn't have to be set out this way, but it may be a useful place to start:



STEP

3 Identify key stakeholders

Confirm the stakeholders that were initially identified as part of your theory of change development. The stakeholders involved with an evaluation may include (but are not limited to!) the team developing and delivering the intervention, those receiving the intervention, academics, research and evaluation team(s), funders, and policy-makers, practitioners, and service providers. Different stakeholders may have different perspectives on what the evaluation should involve or look like, identifying your key stakeholders and engaging them in the evaluation process from beginning can help to ensure the evaluation is shaped around everyone's needs and address any potential barriers.

We've left space below for you to identify your key stakeholders:

STEP

4 Develop evaluation objectives and questions

To ensure your evaluation will help you provide the information you are looking for, it is often useful to consider the following: i) what do you need to understand about the intervention; ii) how do you intend to use the information from your evaluation, iii) what will it help inform and, iv) who will be the intended user of the evaluation.

The aim of an evaluation and evaluation criteria will determine the type of questions we ask, and different evaluation questions will lead to different types of evaluation methods being used and outcomes being measured. As discussed on page 9, the APEASE criteria can be used to structure your evaluation and shape your evaluation questions. The objectives represent more specific statements to help you define your measurable outcomes and steps to meet your aims. Therefore, it's important to clearly establish the aims and objective(s) of the evaluation early on, so that the evaluation questions can be developed accordingly. Key to any successful evaluation in a well-defined evaluation. Even though you may have a number of evaluation questions, it is advisable to make them as specific as possible so that they're answerable.

Behaviourally informed evaluation questions may for example aim to understand the impact of the intervention on changes to the desired behaviour(s), and/or changes to the determinants of that behaviour, or to what extent the behaviour change techniques, or a particular component of the intervention contributed to the desired effect.

Case Study: The Welsh National Exercise Referral Scheme (NERS) is a health intervention that incorporates physical activity sessions and behaviour change content. After an initial assessment of physical activity, health and well-being a 16-week programme of exercise is offered. As a result of the Covid-19 outbreak, face-to-face delivery was suspended, and a remote programme was phased in to replace standard delivery. The University of Hertfordshire and the Public Health Intervention Responsive Studies Team wanted to evaluate the impact of introducing virtual delivery of the NERS programme. Their objectives and evaluation questions are detailed below:

Evaluation Objectives

- How has delivering the Welsh NERS virtually affected whether people join the project, take part in exercise sessions, and complete all 16 weeks of the programme
- Are there things that increase or decrease the chances of people joining and taking part in the Welsh NERS, and what are these for different groups of people?
- How do the outcomes for service users who experience a virtual version of the Welsh NERS compare to those who experience it face-to-face?
- How do the costs of running the Welsh NERS virtually compare to the costs of running it face-to-face?
- What might be the best way to deliver the Welsh NERS in the future and what might be the benefits and downsides of these different options?

Evaluation Questions

- What are stakeholders views and experiences of the NERS when delivered in face-to-face and/or virtual formats, how do they compare and what are the implications for programme uptake, engagement and delivery?
- Does offering a version of the NERS in which some elements may be delivered virtually affect service-user uptake?
- Does implementing different versions of the NERS, in which either some or all elements are delivered virtually affect service-user engagement?
- Does implementing different versions of the NERS, in which either some or all elements are delivered virtually, affect service user retention?
- Does implementing different versions of the NERS, in which either some or all elements are delivered virtually, affect health and wellbeing outcomes?
- What are the expected resources and corresponding costs (including impact on service user out-of-pocket expenses) of delivering core parts of the NERS programme and do they differ for face-to-face and virtual delivery?

You can type in the objective of your evaluation and subsequent research questions below:

The aim of the evaluation is to:

The objective of the evaluation is to:

In order to achieve this, we will ask the following evaluation questions:

-
-
-
-
-

We will use the evaluation findings to help inform:

The intended users of the evaluation are:

STEP 5

Defining your outcome measures

At this stage, it is important to confirm the indicators that were initially identified as part of your theory of change development that will help you measure the changes in your behaviour of interest (outcomes). You can start by confirming your list of possible indicators, based on review of previously published evaluations of similar interventions or commonly used indicators to help measure your behaviour of interest. You should utilise your theory of change here to help prioritise key indicators that can help you test your theory of change or whether the intervention has worked as you intended. At this point, it is also advisable to explore further what data already exist that you can draw upon in your evaluation and in what format, frequency, completeness. That may include, for example, any routinely collected data. You may wish to consider further who holds the data and how it can be accessed. It can also be useful to consider whether you need to also collect any additional data to help answer your evaluation questions.

Choosing measures to assess changes in chosen behavioural outcomes/target behaviour (s) will largely depend on the type of intervention and type of behaviour the intervention is targeting. Where possible, it is best to use objective, valid (accurate) and reliable (replicable) measures of outcomes.

You may be able to utilise already existing data sources or may need to establish a form of data collection (e.g. measure changes in physical activity over a period of time) or track prevalence of the behaviour over time. It is important to also remain consistent with your measures, such as using the same tools to assess any changes in the desired behaviour before, during or after the intervention (so that these can be comparable). In addition to the target behaviour, you can also focus on understanding changes in the determinants of that behaviour (i.e., the barrier(s) that your intervention was aiming to address). For example, you may want to explore if capability, opportunity and/or motivation has changed as a result of the intervention by using validated questions to explore changes to knowledge, understanding, beliefs, and/or attitudes. Identifying means of measuring behaviours, and their determinants, can be challenging so it may be helpful to seek support from subject matter experts, researchers and evaluators, and/or behavioural scientists.

You can answer the following questions to help you define your outcome measures below:

What outcome measures have been used in previous evaluations of the intervention or similar interventions?

What data already exist that I can draw upon for my evaluation?

Who holds the data and how can I get access to this data?

What additional data do I need to collect to help answer my evaluation question(s)?

STEP
6**Select evaluation methodology**

Your evaluation method will be largely determined by the stage the intervention is at, the evaluation questions you are hoping to answer, what data is available and what you may be able to collect, as well as your timelines. Consider whether your evaluation question would be best answered with quantitative or qualitative data, or with a mixed method approach (combining quantitative and qualitative method) which can offer deeper insights. If answering your evaluation question(s) is most suited to qualitative methods, consider who you would need to collect data from, whose perspective do you need to capture (e.g. from service users, individuals who received or were exposed to the intervention, staff who delivered the intervention, policy makers etc). Examples of qualitative data collection methods include: semi-structured interviews, in-depth unstructured interviews, focus groups, observations. Examples of quantitative data collection methods include: surveys, utilising existing structured and unstructured data (e.g. routine health data).

The questions below can be used to help your decision making:

What questions do you hope the evaluation will answer?

What type of methodology will help you answer these questions (qualitative/quantitative)?

What research methods and design type will help you to answer your evaluation questions?

How will you ensure your methodology is as rigorous as possible?

STEP
7**Agree on a timeline and budget**

Confirming a timeline and budget for the evaluation allows us to plan ahead and allocate sufficient resources required to complete the evaluation from start to finish. In reality, your timeframe and budget will often be the key defining factors in what will be possible – i.e., determine the scope and scale of your evaluation.

When reflecting on the above, it is useful to consider if you are going to commission the evaluation externally or internally and associated costs, as well as what resources will be required in order to undertake the evaluation (e.g., consider recruitment, data collection, analysis, reporting and dissemination) and the dedicated amount of time you have.. Once you've identified these, you can establish what may be possible to undertake in terms of the scope and scale of the evaluation to help answer your evaluation question(s).

You can answer the following questions to help you plan, below:

What is your overall budget for the evaluation:

What time do you have to dedicate to the evaluation (e.g. think from commissioning process, designing your evaluation, recruitment, undertaking data collection, analysis, reporting of findings and dissemination):

STEP
8**Select an evaluation team****Independent**

Advantages: An external evaluation team can bring impartiality and credibility into the evaluation. This is because they're objective and have not been involved with the development of the evaluation. The team can bring a particular evaluation expertise and an independent perspective and increase capacity for delivery. Fully independent evaluation teams can be recruited through an external commission.

Disadvantages: Tend to be brought in at the end of intervention for a fixed amount of time, which may limit the ability to be agile and adapt to any unexpected intervention changes; or implement on-going improvements to the intervention. May lack 'insider' knowledge of the intervention, delivery, or context. May be costly, as requires external expertise, time and capacity – so a sufficient budget allocated to the evaluation would be necessary. May also feel pressured to report positive findings by those who have commissioned them.

Another point to consider is – that external commissioning will require management of the evaluation delivery and oversight to ensure it is delivered as per your agreed specifications or contract. This will require allocating additional capacity within your team to oversee the delivery, regular progress meetings with the provider and generally manage the relationship with the commissioned provider.

Semi-Independent

Advantages: It may be possible for another team within the same organisation, who haven't been involved with the intervention, to complete the evaluation. These teams are known as semi-independent and can help to add transparency to the evaluation and offer additional expertise. Can also be cheaper to undertake. **Disadvantages:** May have limited capacity and time.

In-House

Advantages: Evaluations can also be completed by the team who developed and delivered the intervention, although ideally the individual evaluators would not have been directly involved in the intervention development/delivery even if in the same team. 'In-house' evaluations can be cheaper to undertake, and potentially easier to implement on an on-going basis. To help improve the robustness of an in-house evaluation, it can be valuable to engage with external evaluation peer reviewers, or commission outside expertise to help support any particular stages of the evaluation (e.g. data collection). **Disadvantage:** May be perceived as more biased, less reliable and credible. Staff involved may lack capacity or skills to carry out a robust evaluation. Staff may be too involved in the intervention and feel the pressure to deliver positive findings.

What will your evaluation team look like?

Challenges

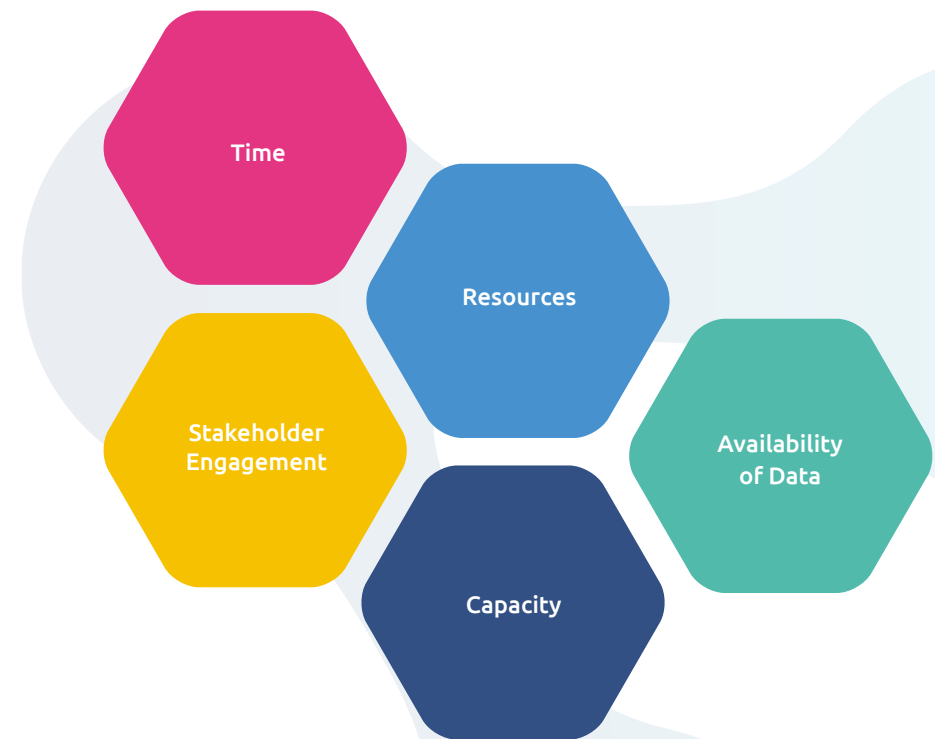
Evaluating the effectiveness of a behaviour change intervention that is being implemented in real-life can present a number of challenges and limitations. Many behaviour change evaluations tend to only consider short-term impact and lack exploration of whether the impact was sustained over time or once the intervention has finished. Often, clearly indicating what is causing the intervention effect may be challenging, as the individual components of the behaviour change intervention often occur together and as such they may also interact together, causing a greater effect or reducing the effectiveness.

Effectiveness may also depend on how the intervention was delivered and in what context or setting which also means generalisation will likely be limited to the specific circumstances. The effect size of behaviour change intervention may often be quite small and opportunity to detect such may depend on having large enough samples for adequate power. The evaluations of a behaviour change intervention may lack a suitable comparison, which may make it harder to attribute with confidence whether the observed impact is due to the intervention or other factors. Very few studies assess the cost-effectiveness of behaviour change interventions or collect any information on cost-effectiveness. Very few studies actually compare impact of the same intervention on different demographic groups or different target groups, which makes it difficult to draw wider conclusions.

To overcome such challenges, it is helpful, where possible, to engage with a range of stakeholders including subject matter experts, behavioural scientists, and researchers and evaluators. In doing so, we can maximise our collective expertise in developing the evidence base for behaviourally informed evaluations and their role in optimising public health outcomes.

Boyce, T., Robertson, R. and Dixon, A., 2008. *Commissioning and behaviour change. Kicking Bad Habits final report* London: The King's Fund. [Link](#)

Michie, S., West, R., Sheals, K. and Godinho, C.A., 2018. *Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Translational behavioral medicine, 8(2), pp.212-224.* [Link](#)



Recap, Summary, Next Steps

This tool aimed to provide an overview of the key points to consider when planning to test, implement and evaluate a behaviour change intervention.

By working through this tool, you have:

- Considered if you should evaluate before, during or after implementation
- Clarified your behavioural specification
- Developed a theory of change
- Identified your key stakeholders
- Developed your evaluation aims, objectives and subsequent evaluation questions
- Defined your outcome measures
- Selected your evaluation methodology
- Noted down your timeline and budget available for the evaluation
- Selected your evaluation delivery team

Now what?

The next step involves the evaluation coming to life; collecting the information you need in order to complete your evaluation and answer the questions that you've outlined in your planning so far.

We've summarised (what we believe to be!) the key take home messages below:

- Behaviourally informed evaluations focus on behaviour change, and/or changes made to behavioural determinants
- Thinking about evaluation from the offset will really help guide the development, implementation, and evaluation of an intervention. It will also enable consideration of the time and resources available for the evaluation and thus a feasible and pragmatic approach
- Identifying behaviour change techniques (the active ingredients) in a scientific and systematic way will help to clarify exactly how your intervention is intending to bring about change
- Developing a theory of change helps to identify what you should be measuring as part of your evaluation and why

A note on dissemination:

It is important to remain transparent with your evaluation and therefore sharing or publishing your evaluation findings can help inform and improve the available evidence base. Wherever possible, we encourage you to share your experience and resources as this not only adds to the evidence base but also enables practice and learning to be shared. The BSU have developed a [case study template](#) which can support you in summarizing, reflecting on, and disseminating your work.

Next Steps:

[BSU Repository](#)

Access all the resources you need from our repository page; curated by us, featuring behavioural science resources from across the globe, including our own.

Public Health Wales Evaluation Community of Practice

For those working in Public Health Wales, the Evaluation Team has developed a suite of evaluation resources as part of the Community of Practice. For more information, you can contact the evaluation team at PHW.Evaluation@wales.nhs.uk

References:

[HMT Magenta Book](#)

[Guide to Evaluating Behaviourally and Culturally Informed Health Interventions in Complex Settings](#)

[Designing and Evaluating Behaviour Change Interventions - The 5 Step Approach to Evaluation](#)

[Recommendations | Guidance | NICE](#)

[Evaluating the Effectiveness of Behavior Change Techniques in Health-Related Behavior: A Scoping Review of Methods Used](#)

[Evaluation of Behavior Change Interventions \(Chapter 22\) - The Handbook of Behavior Change](#)

[Economic Evaluation: Evaluation in Health and Wellbeing - GOV.UK](#)

[RE-AIM – Home – Reach Effectiveness Adoption Implementation Maintenance](#)

[Understanding the Wheel - The Behaviour Change Wheel Book - A Guide To Designing Interventions](#)

[A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance | The BMJ](#)