



Guidance for evaluating interventions preventing household food waste



REFRESH is funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement no. 641933. The contents of this document are the sole responsibility of REFRESH and can in no way be taken to reflect the views of the European Union

Authors

Tom Quested (WRAP)

With thanks to the following reviewers and proof readers for their helpful comments:

Carla Caldeira, Joint Research Centre

Jayne Cox, Brook Lyndhurst

Michael Dorn, University of Bern

Hannah Macey, WRAP

Åsa Stenmarck, IVL

Serenella Sala, Joint Research Centre

Sabrina Stöckli, University of Bern

Erica van Herpen, Wageningen University

Stephanie Wunder, Ecologic Institute

Project coordination and editing provided by Ecologic Institute.

Manuscript completed in April, 2019

This document is available on the Internet at: [optional]

Document title Guidance for evaluating household food waste interventions

Work Package WP3

Document Type Report

Date January 2019

Document Status Draft version 1

Acknowledgments & Disclaimer

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641933.

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of the following information. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

Reproduction and translation for non-commercial purposes are authorised, provided the source is acknowledged and the publisher is given prior notice and sent a copy.

Table of Contents

1	Executive Summary	1
2	Introduction	2
3	Understand the intervention	6
3.1	Logic mapping	6
3.2	Intervention characteristics	9
3.3	Existing knowledge of intervention	13
3.4	Developing evaluation questions	14
4	Develop the evaluation approach	16
4.1	Purpose and audience of the evaluation	16
4.2	Type of evaluation	17
4.3	What to measure, and how?	18
4.3.1	Inputs, resources, activities and outputs	20
4.3.2	Intermediate outcomes	21
4.3.3	Measuring levels of HHFW (final outcome)	23
4.3.4	Alternatives to measuring HHFW using waste compositional analysis	28
4.3.5	Quantifying impacts	29
4.3.6	Understanding the counterfactual	30
4.3.7	Longevity of outcomes and impacts	32
4.4	Data protection	33
4.5	Outputs of the evaluation-development process	33
5	Implement and disseminate the evaluation findings	35
5.1	Implementing the evaluation	35
5.2	Dissemination of findings	35
6	Good practice tips	37

7 Conclusion	38
8 Further reading	39

List of Tables

<i>Table 1: Types of information useful to different audiences</i>	17
<i>Table 2: Examples of metrics for evaluation of household food waste intervention</i>	19
<i>Table 3: Criteria for assessing suitability of metrics / indicators</i>	20

List of Figures

<i>Figure 1: Hypothetical example of a logic model</i>	7
<i>Figure 2: Illustration of levels of household food waste for an intervention and counterfactual</i>	30
<i>Figure 3: Examples of short-lived (left) and long-lived (right) interventions</i>	32

List of Boxes

<i>Box 1: What is meant by an 'intervention'?</i>	3
<i>Box 2: What is evaluation?</i>	4

List of abbreviations

EU	European Union
FUSIONS	Food Use for Social Innovation by Optimising Waste Prevention Strategies
HHFW	Household food waste. In this guidance, HHFW is used to refer to the sum of wasted food (the edible parts of items) and inedible parts associated with food (such as bones, egg shells and inedible rinds of fruits).
NGO	Non-governmental organisation
RCT	Randomised Control Trial
REFRESH	Resource Efficient Food and dRink for the Entire Supply cHain
WP	Work Package
WRAP	Waste & Resources Action Programme

1 Executive Summary

This document provides guidance for evaluating interventions and policies designed to reduce the amount of household food waste (HHFW). In this context, interventions are any activity – such as campaigns, changes to food packaging or products – that are being undertaken in order to prevent HHFW.

At the time of writing, there are very few evaluations relating to HHFW-prevention interventions, meaning that it is not known what the best ways are to reduce the amount of HHFW in different contexts. This guidance aims to address this evidence gap.

The guidance is for anyone undertaking or studying interventions who would like to understand whether the intervention has been effective in preventing HHFW. Use of the guidance could also help understand why a particular intervention has been successful (or not, as the case may be). Applying this guidance will also produce information useful to undertaking studies calculating the efficiency of policies and interventions: comparing the resources used to the level of reduction in HHFW. However, guidance on these efficiency studies are not covered in detail in this document, although references are provided.

The guidance is designed to provide a practical overview of the subject, covering the key steps of an evaluation:

- Understanding the intervention in question (Chapter 3)
- Designing the evaluation appropriately (Chapter 4)
- Implementing and disseminating the evaluation (Chapter 5)

Links are provided to a range of useful resources: general evaluation guidance, as well as material specific to HHFW. A summary of good-practice tips is also provided.

In the future, use of this guidance should lead to a step-change in the quality of studies evaluating HHFW. This should provide evidence for policy makers and other decision makers to select the most appropriate approaches, so that they are able to reduce the amount of food wasted from households in a cost-effective manner.

2 Introduction

Food waste has emerged as an important global issue, having substantial negative impacts. In high-income countries, food waste from households is the single-largest contributor to these impacts¹. Over the last decade, there have been many efforts to help households waste less food; however, few have been evaluated with sufficient detail to allow determination of which types of interventions are most effective.

This document is designed to help address this lack of evidence. It provides guidance on how to evaluate the effectiveness of 'interventions' designed to reduce the amount of household food waste (HHFW). This guidance has been developed as part of the EU Horizon 2020 REFRESH and is closely linked to the policy brief and supporting documentation on consumer food waste².

The guidance reflects the steps required to complete a good evaluation. Chapter 3 covers the steps required to understand the intervention itself, an important part of the evaluation process. Chapter 4 discusses the development of an evaluation plan. This includes considering the purpose and audience for the evaluation, the type of evaluation to choose and what information to gather as part of the evaluation. Chapter 5 includes implementation and dissemination of the evaluation findings, while Chapter 6 contains some good-practice tips.

Who is this guidance for? This guidance is designed for four groups of people:

- Policy makers in governments who are working to prevent HHFW, including regional, national and local governments.
- Other decision makers and people running HHFW prevention interventions (e.g. charities and NGOs, food retailers, community groups)
- Evaluation specialists who are involved in studies relating to HHFW prevention
- Academic researchers whose research focuses on identifying and testing solutions to prevent HHFW

¹ For example, see Stenmarck et al. (2016), *Estimates of European food waste levels*, FUSIONS report

² Wunder, 2019. *REFRESH Policy Brief: Reducing consumer food waste*. EU Horizon 2020 REFRESH. <https://eu-refresh.org/node/908/>
Wunder et al. (2019). *Policies against consumer food waste. Policy options for behaviour change including public campaigns*. EU Horizon 2020 REFRESH <https://eu-refresh.org/policies-against-consumer-food-waste>

In general, it is good practice for those implementing interventions (the first two groups indicated above) to work collaboratively with people from the latter two groups, who are well-positioned to study the effectiveness of interventions.

Many of the processes and principles described in this guidance are relevant to interventions to prevent food waste in other sectors (e.g. restaurants and canteens), but many of the measurement methods describe in this guidance would not be appropriate. We therefore focus on evaluation of household food waste, rather than including guidance relating to food waste from consumption out of the home.

Box 1: What is meant by an 'intervention'?

For the purposes of this guidance, 'intervention' is used as a broad term: any activity that has been designed to reduce the amount of HHFW. This could include activities by:

- Government (e.g. legislating on date labels)
- Businesses in the supply chain (e.g. changing the size of packs to help people buy the right amount of food for their needs), or
- Charities and community groups (e.g. running a HHFW prevention campaign).

Sometimes a single intervention is evaluated. In other circumstances, a range of interventions will have been designed to work together to prevent HHFW and these can be evaluated together.

Why is evaluation important for HHFW prevention interventions?

Evaluation helps a range of people to understand whether an intervention has produced the desired outcomes. It helps to understand how the desired outcomes have been achieved, and the types of people who have been influenced. It also allows the cost-efficiency of an intervention to be assessed. This helps with important decisions relating to the interventions:

- For **new interventions**, previous evaluations can help decide what type of intervention to use, and provide information useful to the design process
- For **existing interventions**, evaluation can help decide whether an intervention needs stopping if it is ineffectual, altering to increase its effectiveness, or rolling out more widely if it is working well

Box 2: What is evaluation?

Evaluation is a process to understand an 'intervention': how it was implemented, the effects that it had, for whom, how and why.

An evaluation examines how an intervention was designed, carried out and what the results of the intervention were. Therefore, evaluations are practically-focused: they investigate what actually happened in practice, rather than what was expected.

Evaluations use a range of analytical methods to collect and analyse information. The specific methods employed depends on a wide range of factors; methods appropriate to interventions focusing on HHFW prevention are discussed in this guidance.

At the time of writing, there is a lack of evidence about how effective different interventions are at preventing HHFW (see further reading for references outlining this problem). This makes it difficult for policy makers and other decision makers to make evidence-based decisions.

By developing clear and consistent guidance, this problem can be resolved: a greater number of appropriate evaluation studies can be conducted. These will have greater consistency, allowing comparison of results, helping to determine which types of evaluation work best in different circumstances. This will allow resources allocated to preventing HHFW to be spent effectively.

When should I be considering evaluation? If you are in the process of developing an 'intervention' to reduce HHFW, then the time is right to design the evaluation. Generally, the development of the intervention and its evaluation should progress in parallel. All too often, evaluation is only considered towards the end of the implementation phase, which is usually far too late for effective evaluation.

If the planning of an evaluation is started too late, it is often not possible to access all the data and other information required to understand whether the intervention was successful. For instance, a delay in evaluation planning may impede collecting data prior to the start of the intervention, making it difficult to track changes over time. It also means that key decisions about the evaluation are rushed, which usually results in a poor-quality evaluation that doesn't help policy makers and other decision makers.

Who can help me evaluate? If you are developing an intervention to reduce HHFW (e.g. a government official developing policy, a retailer changing how products are priced, packaged and sold), it is important to collaborate with evaluation practitioners, as well as researchers who have knowledge of HHFW prevention. They can help you through the process described in this document, navigating some of the difficult decisions that will need to be

made. Furthermore, there are many resources available to help develop an evaluation³.

³ For example: <https://betterevaluation.org/> and <https://www.evaluation.org.uk/index.php/news-resources/ukes-publications/46-ukes-guidelines-for-good-practice-in-evaluation>

3 Understand the intervention

Before the evaluation can be designed, there are a few steps that need to be undertaken to ensure the intervention is well understood. These involve:

- Logic mapping (Section 3.1)
- Characteristics of the intervention (Section 3.2)
- Determining existing knowledge of the intervention (Section 3.3)

These issues are explored in the sections below. It is important that these stages are undertaken because household food waste is a complex issue, due to it being the result of many inter-related activities⁴.

3.1 Logic mapping

A first step in developing an evaluation is to set out:

- The design of the intervention: what will be done, by whom, when, etc.
- The desired outcomes and impact of the intervention
- The resources available for the intervention, and
- How these above points relate to each other

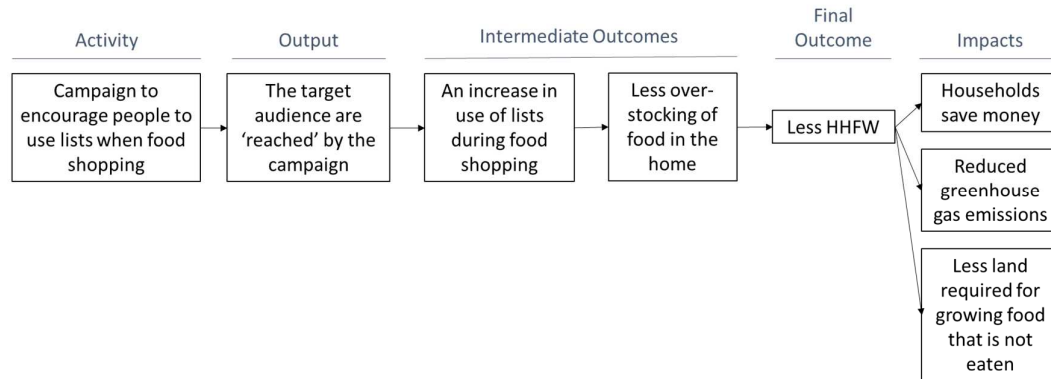
A commonly used method to support this step is to develop a 'logic model', also known as 'programme theory' or 'intervention logic'. A logic model describes the 'theory' of the intervention: i.e. how the activities of the intervention will lead to the 'final outcome' via the outputs and outcomes. The benefits of logic mapping are that it:

- Provides a clear and shared understanding of how the intervention is designed to work
- Helps identify assumptions inherent in the intervention
- Provides a framework for viewing existing evidence associated with the intervention (and the intervention's assumptions)
- Provides an early check as to whether the intervention is likely to achieve its impact, and
- Helps identify factors to measure (or qualitatively assess) as part of the evaluation

⁴ See [Quested et al. 2013, *Spaghetti soup: The complex world of food waste behaviours, Resources, Conservation and Recycling*, 79, 43-51](#) and Wunder et al. (2019). Policies against consumer food waste. Policy options for behaviour change including public campaigns. EU Horizon 2020 REFRESH <https://eu-refresh.org/policies-against-consumer-food-waste>

An example of a simple, hypothetical logic model for a HHFW prevention intervention is shown in Figure 1.

Figure 1: Hypothetical example of a logic model



The list below provides definitions of each stage of a logic map (see Table 2 for more details):

- **Activities:** What was done in the intervention, e.g. radio commercials developed, changes to the packaging of a range of food items
- **Outputs:** The direct results of the activities, e.g. the number of people reached by a campaign, the newly designed packaging being available in grocery stores.
- **Intermediate outcomes:** Changes resulting from the intervention that allow the final outcome and impacts to be achieved
- **Final outcome:** The focus of the intervention: in this case, reducing the amount of HHFW
- **Impacts:** Changes that occur as a result of the final outcome, for example environmental or social impacts

In addition, **inputs and resources** should be recorded (i.e. the resources available for intervention such as money and people’s time). This is especially important if cost-efficiency is to be assessed as part of the intervention.

The logic model provides a framework for discussing the assumptions and evidence relating to the intervention. For each arrow in the logic model, there will be a set of assumptions; if these assumptions are not met in the real world, the intervention may not be effective.

For instance, for the arrow between the output (‘The target audience are reached...’) and the first intermediate outcome (‘An increase in the use of lists...’), there is an assumption that the campaign is effective at changing this activity for the target audience. There may already be evidence to support this assumption, e.g. via a pilot study that tested campaign materials with a small group of the target population in realistic conditions, or from

previous research. However, in many situations, there will be limited or no evidence about an assumption. In such cases, testing this assumption through the design of the evaluation should be considered (see below). This helps to identify what would be useful to measure during the evaluation.

Some assumptions may relate to other activities: for instance, in the hypothetical example in figure 1, the campaign may raise awareness of the need to use shopping lists, but the evidence may suggest that this will only be successful if there is also, say, prompting by supermarkets of their customers (i.e. supermarkets providing a reminder to their customers at an appropriate time). If a 'prompting' intervention is also being considered, this could be added to the logic map so – in this case – two interventions are shown aiming to achieve the same outcomes and impacts.

The logic mapping can provide an early check as to whether the intervention is likely to achieve its stated outcomes and impacts. It is relatively common for logic mapping to trigger a re-design of the intervention such that it is more likely to be effective.

Logic mapping should be conducted with those responsible for implementing the intervention, those conducting the evaluation and other key decision makers associated with the intervention (e.g. the budget holder). By creating an agreed, clear framework, subsequent steps of the evaluation are easier to conduct.

This guidance document does not present a full 'how-to' manual for logic mapping. For further information on how to perform logic mapping see *The Magenta Book*⁵, and *Logic Mapping: Hints and Tips*⁶.

In the design of the intervention, a range of techniques are likely to be used (e.g. stakeholder mapping, brainstorming). Although this guidance does not discuss methods primarily focused on the design of the *intervention*, the outputs from these techniques will also be useful for designing the *evaluation*. Therefore, there is considerable benefit in sharing a wide range of information about the intervention with those performing the evaluation, as it will help the evaluators understand the intervention better, leading to a more appropriately designed evaluation.

⁵ HM Treasury 2011, *The Magenta Book*, Guidance for evaluation: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

⁶ Tavistock Institute (for the Department of Transport), *2010 Logic Mapping: Hints and Tips Guide*: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/3817/logicmapping.pdf

3.2 Intervention characteristics

The nature of the intervention will influence the design of the evaluation. This section looks at some different factors that are important to evaluation.

What is the intervention seeking to influence?

This guidance covers interventions seeking to prevent household food waste (HHFW). However, different interventions may aim to influence different 'fractions':

- All HHFW⁷
- Just the wasted food (edible parts)
- Only one food category (e.g. bakery; for example, a campaign that focuses on hints and tips solely on this category)
- Only a single product (e.g. one particular product from a manufacturer; for example, where a change has been made to the product itself or the packaging with the aim of stopping it from being wasted)

This has implications for what gets measured: when monitoring the levels of food waste (section 4.3.3), the definition used should reflect the aims of the intervention: for instance, if any intervention is likely only to impact on bakery, then the measurement of HHFW could focus solely on that food category. However, it is useful to consider whether an intervention could have wider impacts than its primary focus: for instance, a campaign focused on bakery waste could also lead to a reduction in the amount of other products that get wasted (i.e. a 'spill-over' effect).

In addition, if an intervention is aimed at a single product, then this will influence how monitoring information is gathered. When assessing intermediate outcomes via surveys, it may be necessary to target the survey only to people who purchase that product, rather than at the population as a whole.

Who is the intervention seeking to influence?

To design an evaluation, it is also important to understand who the intervention is seeking to influence (the audience). For example, is the audience only the people who have engaged with the intervention, e.g. all those who have engaged with a campaign? Or is it everyone in a geographic area where the campaign was focused? Answering this type of question helps

⁷ In this guidance, household food waste is used to refer to the sum of wasted food (the edible parts of items) and inedible parts associated with food (such as bones, egg shells and inedible rinds of fruits).

us to understand how 'evaluable'⁸ the intervention is. In the former case, evaluation would require identifying and contacting those who have engaged with the campaign (see also comments on GDPR, below). This may present challenges in creating a sample with which evaluation-based research can be undertaken. However, if there is an effect of the campaign, it may be easier to identify than sampling from the overall geographical area, where any effects (e.g. less HHFW) are likely to be more 'dilute'.

It is also helpful to understand the attributes of the audience: their existing capabilities, knowledge, resources, etc. A common finding from evaluations is that interventions often demonstrate little or no impact if the context of the audience has not been well understood. Considering this point allows assessment of the intervention mechanisms: are they appropriate for the characteristics of the audience, based on existing evidence?

For evaluation purposes, knowing the audience starting point of the audience can also help interpret the effects observed towards the end of the evaluation. For example, what proportion of them are already undertaking the desired behaviour? What is the starting amount of HHFW? How much room is there for change?

What type of intervention is used?

Some interventions will focus on engaging people to do something different. This could be via one or more of the following⁹:

- **Provision of information**, e.g. on the impacts of food waste
- **Prompting** people to undertake a 'desired' behaviour¹⁰, e.g. providing reminders at the appropriate time for storing leftovers
- **Modelling behaviour**, i.e. demonstrating a 'desired' behaviour, for example via on-line videos showing people how to do a particular activity
- **Commitment**: asking people to commit to undertake a 'desired' behaviour, e.g. via a public pledge
- **Providing feedback** on the behaviour – i.e. information on whether the behaviour has been successfully performed
- **Rewarding** 'desired' behaviour, e.g. through some form of prize or payment

⁸ 'Evaluable' is the degree to which the intervention can be practically evaluated

⁹ Adapted from Stöckli, Niklaus, Dorn, 2018: *Call for testing interventions to prevent consumer food waste*, Resources, Conservation and Recycling, 136, 445-462

¹⁰ In this context, 'desired behaviours' are activities that can help reduce the amount of food waste produced. These include planning meals, making shopping lists, storing food optimally, cooking appropriate amounts, storing and using leftovers.

- **Penalising** 'undesired' behaviour, i.e. the opposite of rewarding
- **Building capability:** supporting the development of skills and confidence with regard to the 'desired' behaviours

Other useful classifications and resources relating to behaviour change exist. For example, MINDSPACE¹¹, EAST¹² and the ISM model¹³ have been developed to support the design of interventions and policies.

In contrast, there are interventions that attempt to prevent HHFW by making changes to food products, their packaging and how they are sold. These could include changes to:

- Size of food packs available in shops
- Relative prices of foods, e.g. such that smaller packs are a similar price per kilogramme compared to larger packs
- Promotions on food items, so that people are not encouraged to buy more than they need
- Shelf life of food items
- Packaging: e.g. functionality that allows half the pack to be opened, but the other half to remain under a modified atmosphere.

A further category of intervention seeks to influence the environments that are relevant to food becoming wasted. These include:

- Provision of kitchen gadgets that support 'desired' behaviours: e.g. measuring devices, storage containers
- Infrastructure that facilitates food-related activities: e.g. cameras in the fridge, apps that integrate meal planning and purchasing foods

Finally, some interventions will be designed to provide a 'framework' to encourage a range of the above to be enacted. This could include changes to legislation, taxation, other financial incentives.

The above lists are not meant to be exhaustive and there will be overlaps between the groups. However, it indicates that there is a wide range of interventions that will require different evaluation approaches. The outputs of the intervention will be different between these groups, as will many of the intermediate outcomes. The time lag between implementation of the intervention and the outcome on food waste may also differ significantly

¹¹ <https://www.instituteforgovernment.org.uk/our-work/policy-making/mindspace-behavioural-economics>

¹² <https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights/>

¹³ <https://www.gov.scot/publications/influencing-behaviours-moving-beyond-individual-user-guide-ism-tool/>

between these groups. All these factors need to be considered when developing the evaluation plan.

Do people have to consciously engage with the intervention?

Closely linked to the previous discussion, some interventions are designed to engage people, e.g. campaigns that raise awareness of the issue of food waste or prompt people to change a particular behaviour linked to food waste. However, others seek to influence the level of food waste without people necessarily being aware of the change: increasing the shelf life of a product or using behavioural economics to 'nudge' people, e.g. by providing a subconscious cue¹⁴.

This distinction is important for evaluation: in the first, it may be important to find out how many people can recall a campaign or information provided to them. However, in the second group this type of information is unlikely to be important to the evaluation.

In addition, the first type of intervention is likely to influence many of the potential measurement techniques. If a campaign is successful in increasing a population's awareness of food waste, then it may become less socially acceptable to waste food. This could lead to people answering questions in surveys or completing food-waste diaries in a way that is consistent with this change. A good evaluation will minimise the impact of this issue through carefully designed monitoring and analysis.

Single intervention or a programme of multiple interventions?

An important distinction is whether the evaluation focuses on a single intervention or a programme containing multiple, complementary interventions. If there are multiple interventions, it is important to understand at what point in the logic map these interventions 'converge': for example, do they all aim to encourage list making, or do they each aim to influence a different behaviour related to food waste (and therefore converge at the final outcome stage (less HHFW) of the logic map). This will help decide which are the key elements of the logic map, and therefore which metrics are most important to monitor.

Where an intervention aims to influence behaviours, it is important to define these clearly. This allows the evaluation to proceed more easily. This is not straight-forward for interventions focusing on HHFW, because HHFW can be influenced by a large number of interacting behaviours¹⁵. For example, to help households buy the right amount of food, an intervention may focus on shopping-list making; however, this may only be effective in reducing HHFW

¹⁴ For examples of nudges, see:

<https://www.behavioraleconomics.com/resources/introduction-behavioral-economics/>

¹⁵ As described in [Quested et al. 2013, *Spaghetti soup: The complex world of food waste behaviours*, Resources, Conservation and Recycling, 79, 43-51.](#)

if it also accompanied by people planning some of their meals and checking the food they have in the fridge and cupboards. The interaction between the behaviours may be important to the outcome. Evaluators and those running interventions may need to spend some time determining which are the important behaviours and their interactions and how the intervention will influence them. These can be captured on the logic map of the intervention (see section 3.1).

3.3 Existing knowledge of intervention

Prior to implementing an intervention, it is necessary to review what is already known about the intervention in similar contexts to your own. This can help the design of your intervention and its evaluation. This should include searches of the academic literature and the grey literature¹⁶. As noted by Stöckli et al., there is much information in a range of fields: for instance, consumer and environmental psychology. In addition, we can learn from food-related behaviour change (e.g. relating to health), behavioural economics and sociological studies.

If there is a lack of evidence about whether interventions of the type being considered are effective – which, at the time of writing is likely to be the case – then it would be wise to consider evaluation / research that can help fill this gap. This could involve one or more of the following:

- **Testing in controlled conditions:** this would involve testing the intervention in controlled conditions¹⁷, such that its effectiveness can be assessed. These controlled conditions will be designed so that a scientific comparison can be made: e.g. between a control group and an intervention group. Nonetheless, the conditions should be as realistic as possible, otherwise the effectiveness of the intervention may be different compared to the real world. This type of study is likely to involve academic researchers working alongside other interested groups. It could provide information useful to the following types of evaluation: process, empirical impact, theory based (see Section 4.2); if costs of the intervention are known, empirical impact data can be used for economic evaluation.

¹⁶ The following two papers provide a summary of relevant literature in the academic literature, and some information from the grey literature:

Reynolds et al. 2019, Consumption-stage food waste reduction interventions – what works and how to design better interventions, *Food Policy*, 83, 7-27

Stöckli, Niklaus, Dorn, 2018: Call for testing interventions to prevent consumer food waste, *Resources, Conservation and Recycling*, 136, 445-462

¹⁷ e.g. in a simulated situation such as the hypothetical scenario in Holthuysen et al. (2016), *The effect of date marking terminology of products with a long shelf life on food discarding behaviour of consumers*: <http://edepot.wur.nl/428726>

- **User experience:** this involves understanding how people interact with the intervention (e.g. with campaign material or technology designed to prevent household food waste). It can be performed on a fledgling idea for an intervention, the fully designed intervention or anywhere in between. It can involve a range of methods: observation of people's interaction with the intervention, interviews with testers and data relating to the intervention (e.g. amounts of food waste generated in tester's home). This type of study is likely to involve user-experience practitioners working alongside other interested groups. It usually focuses on process evaluation (did the intervention work as intended?) but can also gather other information.
- **Piloting with evaluation:** this involves deploying the intervention on a small scale. This has the advantage of understanding the strengths and weaknesses of the intervention in real situations, at a lower cost than implementing at full scale. The type of evaluation used alongside the pilot would depend on other factors (see in Section 4.2).

3.4 Developing evaluation questions

One key output of this part of the evaluation process is to develop the research questions that the evaluation seeks to answer, often referred to as the evaluation questions. These questions should be a key focus of the development of the evaluation¹⁸. They emerge from a number of different activities:

- Logic mapping
- Discussions with policy makers and decision makers about what they would like the evaluation to answer
- Assessment of what is already known of the intervention (the evaluation questions usually focus on areas with less existing knowledge)

Evaluation questions can generally be classified into two sorts:

- Process questions, which focus on the implementation of the policy or intervention. This usually link to the inputs, activities and outputs recorded in the logic map.
- Outcome (or impact) questions, which are concerned with the effects of the intervention or policy.

In many evaluations, it is not possible to answer all the evaluation questions due to the resources available. Therefore, questions should be prioritised, considering how important it is to answer each question and the resources

¹⁸ In addition to general evaluation references, the following guidance is useful for developing evaluation questions:

<https://www.albertahealthservices.ca/assets/info/res/mhr/if-res-mhr-eval-resources-plan-framework.pdf>

required to answer each. This prioritisation exercise is an important part of developing the evaluation approach (see Chapter 4).

This section has described the process for understanding the intervention. This allows all those involved in the evaluation to develop an evaluation plan, the subject of the next chapter.

4 Develop the evaluation approach

Once the intervention is understood (Chapter 3), an **evaluation plan** can be developed. This should cover the objectives of the evaluation, the research questions that the evaluation attempts to answer, how the evaluation will be performed, the timescales and how the findings will be disseminated and used by different groups. In addition, the plan should include the outputs of the evaluation, their timing and any other milestones.

The evaluation plan helps to facilitate engagement with the range of stakeholders. It also allows the issues and risks related to the evaluation to be identified early in the evaluation process, helping to mitigate and manage them. The evaluation plan can also be reviewed by relevant practitioners at an early stage to assess whether the planned evaluation will be likely to provide robust evidence; corrective action can be taken if necessary.

This chapter covers various aspects of developing an evaluation plan:

- Section 4.1 covers the purpose and audience of the evaluation.
- Section 4.2 discusses the range of evaluation types and how each develops an understanding of different aspects of an intervention. This section also provides guidance on how to select the appropriate type for your purpose and audience.
- Section 4.3 focuses on what to measure, and how to measure it. It discusses the range of metrics from across the logic map and how to select the appropriate ones. It also contains advice on how to measure the 'counterfactual' (what would have happened in the absence of the intervention) and longevity effects.

4.1 Purpose and audience of the evaluation

While designing the evaluation, it is useful to consider who the main users of the findings will be and how they will be engaged. Findings from evaluation of HHFW prevention interventions could be useful to many different people, including those:

- Implementing the interventions, helping to improve the intervention
- Assessing the value of the intervention (e.g. auditors, parliament, the media)
- Making decision about this (or future) interventions, including budget holders
- Making funding applications to run similar interventions
- Engaged in research on this topic

It is likely that the audience for your evaluation will include most of, if not all, these groups of people. Given that resources are usually constrained for interventions and evaluation, it is important to prioritise each of these groups.

For the priority audience(s), it is important to consider the type of information that these groups would like from an evaluation. Table 1 provides some common requirements for different audiences.

Table 1: Types of information useful to different audiences

Evaluation audience	Types of information
Intervention implementers	Information on which elements of the intervention are working well (or not)
Funders, budget holders, auditors	Demonstration of activity and output Estimation of outcomes, impact and cost efficiency of the intervention
Researchers	Consistent information allow comparison between studies Information that allows links in logic map to be assessed

4.2 Type of evaluation

The type of evaluation used will depend on the logic of the intervention (Section 3.1), its nature (3.2), what's already known about interventions of this type (3.3), why the evaluation is being conducted and for whom (4.1).

There are several of types of evaluation. Some (*ex-ante*) focus on estimating the impact of an intervention *before* it is implemented. Others (*ex-post*) focus on understanding and / or quantifying the impact of an intervention *during or after* its implementation. It is this latter category that this guidance focuses upon. There are a range of types of *ex-post* evaluations:

- **Process evaluation:** this explores how an intervention was implemented (e.g. discovering if it was implemented as initially intended). This uses information from individuals involved or affected by the intervention to assess its effectiveness.
- **Empirical impact evaluation:** this type of evaluation tests whether an intervention was associated with any significant changes in the outcomes of interest. These use quantitative data that has been measured during or after implementation of the intervention.
- **Theory-based evaluation:** these involve increasing the understanding of the connection (i.e. the theory) between an intervention and the anticipated impacts. These connections can be explored using a wide range of research methods (both qualitative and quantitative).

- **Economic evaluation:** these calculate the economic costs associated with an intervention and compare these to the benefits, translated into economic terms. This type of evaluation usually requires information derived from process and empirical-impact evaluation: the economic benefits associated with HHFW prevention will depend on the degree to which the intervention was successful in preventing HHFW.

For more details on these evaluation types, see the Magenta Book¹⁹ as well as references for specific types of evaluation²⁰.

It is common for evaluations to incorporate elements from multiple evaluation types, as they are often being undertaken for a range of reasons. The choice of evaluation method(s) should reflect the evaluation questions. For instance, if how the intervention works is poorly understood, then elements of process and theory-based evaluation can be useful. If determining the degree to which the intervention reduces HHFW is a priority, then empirical impact evaluation is important.

Decisions on which type of evaluation to undertaken should be taken in consultation with evaluation practitioners.

4.3 What to measure, and how?

What to measure, and how to measure it, is another key decision for those evaluating HHFW prevention interventions. Most evaluations will need to capture a range of metrics. All metrics chosen should have a clear link to the evaluation questions (section 3.4): they should provide information that helps to answer one of these questions.

The logic map is invaluable in guiding decisions on important quantities to measure. This will highlight important points in the logic: candidates for evaluation metrics include quantities important to the intervention and areas where there is much uncertainty about the logic of the intervention. The potential types of metric are summarised in Table 2, classified by the stages of the logic map.

¹⁹ HM Treasury 2011, The Magenta Book, Guidance for evaluation: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

²⁰ For example, theory-based evaluation: <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/theory-based-approaches-evaluation-concepts-practices.html>

Impact evaluation: <https://www.gov.uk/dfid-research-outputs/impact-evaluation-a-guide-for-commissioners-and-managers>

An example of an economic evaluation in the area of food waste is: <https://champions123.org/the-business-case-for-reducing-food-loss-and-waste/>

Table 2: Examples of metrics for evaluation of household food waste intervention

Metric type	Description	Examples	Examples of how to measure
Inputs and resources	Resources available for intervention; important for economic evaluations	Budget (money) used; people's time; in-kind support (e.g. use of facilities, materials)	Keeping consistent records during intervention
Activities	What was done in the intervention	Campaign 'materials' (e.g. leaflets, adverts); new packaging technology; change to formulation of food product	Keeping record of new materials, packaging, etc.
Outputs	The direct results of the activities	No. of people reached by campaign; number of training sessions run; number of reformulated products sold	Collating relevant info: sales data; communications metrics
Intermediate outcomes	Changes resulting from the intervention that allow the final outcome and impacts to be achieved	Changes in awareness, attitudes, knowledge, skills, behaviour (e.g. storing leftovers, planning meals)	Use of questionnaire surveys and interviews
Final outcome	The focus of the intervention	Change in the amount of HHFW	See section below
Impacts	Changes that occur as a result of the final outcome	Changes in greenhouse gas emissions, water use, agricultural land requirements, fertiliser use, etc.	Modelling (see below)

To help guide decisions around which metrics / indicators to include in the evaluation, criteria have been developed to help decide on the quantities to measure (Table 3)²¹.

²¹ Burgos et al. (2016) Policy Evaluation Framework, FUSIONS report: <http://www.eu-fusions.org/index.php/download?download=255:policy-evaluation-framework>

Table 3: Criteria for assessing suitability of metrics / indicators

Criteria	Definition
Attainable	The measurement of the indicators should be achievable by the policy or project and should be sensitive to the improvements the project/policy wishes to achieve.
Clear	Indicators should effectively target the factor which they are measuring and should avoid ambiguity and arbitrariness in the measurement.
Comparable	The indicator measurement should enable comparison over the different life-cycle stages of the policy or project, as well as between different policies or projects.
Comprehensible	The definition and expression of the indicator should be intuitively and easily comprehensible to users.
Cost-effective	The cost of collecting and processing the data needed for the chosen indicators should be reasonable and affordable.
Up to date	Indicator information should be as up to date as possible, to reflect current or recent circumstances. The impact of delays between collection and use should be considered and factored into the analysis, where necessary using extrapolation techniques.
Measurable	Indicators should be defined so that their measurement and interpretation are as unambiguous as possible, preferably using data that is readily available, relevant, reliable and meaningful.
Redundant	While each input variable should measure a discrete phenomenon, separate indicators that measure the same phenomenon may be necessary and desirable.
Relevant	Indicators should be directly relevant to the issue being monitored or assessed and should be based on clearly understood linkages between the indicator and the phenomena under consideration.
Reliable	The results from an indicator should be replicable by different researchers using standard methods. The methods should be stable over time and as valid in as wide a circumstance as possible.
Sensitive	Indicators should be able to reflect small changes in those things that the actions intend to change.

More guidance is given for metrics relating to each area of the logic map in the following sub-sections. Further guidance in this area will be published later in 2019²².

4.3.1 Inputs, resources, activities and outputs

Measurement of metrics in this area often involves keeping accurate records of relevant quantities during the intervention. This will help in understanding

the effectiveness of the intervention once the information has been collated and analysed.

Once the relevant input, resource, activity and output metrics have been decided, it is important to ensure that there is a process for capturing this information to maximise the amount of relevant data collect. Ensuring that all relevant people are informed of these recording requirements is important: if there is a need to capture the number of people attending a training course, those running the courses need to be informed of this need²³. Despite this being a seemingly obvious point, it is often overlooked.

For some metrics, the data may be difficult to obtain or more sensitive. A common example relevant to HHFW is sales data relating to food items (e.g. sales of items with new packaging). Sales data could come from the manufacturers and retailers involved with the production and sale of the product. However, they may view this data as sensitive. It is best to discuss this with those who own the data during the evaluation development: there may be ways of handling or publishing the data that remove the sensitivity (e.g. publishing data for a range of products in aggregate (total), rather than providing data for individual products). Alternative sources of sales data are companies that run panels of people who scan in their food purchases²⁴.

More information on obtaining relevant metrics of communications campaigns can be found in the Evaluation Framework 2.0, Government Communication Service²⁵. Food-waste examples of metrics can be found in a FUSIONS report on policy evaluation²⁶.

4.3.2 Intermediate outcomes

The intermediate outcomes that are important for a HHFW prevention intervention can vary widely, depending on the nature of the intervention.

²² Caldeira, De Laurentiis, Sala, Assessment of food waste prevention actions: development of an evaluation framework to assess performance of food waste prevention actions, Interim document - Limited distribution, 2019, JRC115987 (due to be published later in 2019)

²³ Informing those involved in the intervention is necessary, but often insufficient to ensure people collect the appropriate information. Everyone involved in the intervention needs to be convinced of the value of evaluation so that they prioritise this data collection – this includes the organisations delivering the activity. In some settings, delivery organisations might see evaluation as a distraction from delivering the intervention itself and may not participate in the necessary data collection.

²⁴ e.g. Kantar WorldPanel or Nielsen Scantrack

²⁵ <https://gcs.civilservice.gov.uk/guidance/evaluation/tools-and-resources/>

²⁶ Burgos et al. (2016) Policy Evaluation Framework, FUSIONS report: <http://www.eu-fusions.org/index.php/download?download=255:policy-evaluation-framework>

This is where the logic map is especially helpful in guiding the choice of metrics.

The ideal method for determining whether there has been any impact on intermediate outcomes is through direct observation. For example, if the intervention aims to influence where people store certain food items (e.g. to optimise their shelf life), one method would be to check this in a representative sample of people's homes (with permission).

However, direct observation can be expensive and can be difficult to observe for some quantities: it is hard to assess whether skills or confidence in certain food-related tasks have increased in an objective way. Behaviours (such as meal planning) can easily be influenced by the observation itself.

In many cases, a cost-efficient solution to monitoring intermediate outcomes is to use questionnaire surveys. These allow a range of questions to be asked about multiple metrics. The sample can be tailored to reflect the intervention. For example, if a nation-wide campaign is being evaluated, the sample can be nationally representative. At the other extreme, if the intervention focuses on providing training to a small group of people, the questionnaire can be deployed to some (or all) of the training participants.

Particular attention should be given to minimise bias in the questionnaire. The wording of the questions, the response options, the order of the questions and how the questionnaire is described to the participants could all influence the results. Good questions should:

- Be neutral, avoiding any implied criticism or indication of what the 'correct' answer is
- Be as objective as possible: be clear on what you are asking, avoiding ambiguity
- Ask about one thing at a time, otherwise the question may be difficult to answer
- Minimise social desirability bias²⁷, e.g. by only mention food waste where it is critical to that particular question
- Provide a range of response options that cover most eventualities, including – where applicable – 'none of the above', 'don't know' and 'would prefer not to say'.

Although designing survey questions can appear simple and there is much published advice²⁸, it is recommended that someone trained in developing questionnaires is consulted during this step.

²⁷ 'Social desirability bias' is where people respond to surveys in a way which they think is viewed favourably by others.

²⁸ e.g. Boynton, P.M., Greenhalgh, T. (2004) Selecting, designing, and developing your questionnaire, *BMJ*, 328, 1312–1315

Another useful step in developing questions is cognitive testing²⁹, where resources allow. Cognitive testing describes a range of methods which capture the respondents' thought processes and understanding of survey questions. It often involves asking the survey questions to a small sample of people and then asking them why they responded in the way that they did. This is useful because it can highlight any problems research participants can have in answering questions (e.g. words that are not commonly understood). This is particularly important for questions relating to food waste because the language used to talk about food waste can have a strong influence on how people answer a question. In addition, food waste can be an emotive topic for respondents so finding the right language to elicit an honest response is important.

There is a growing amount of existing research that includes survey questions relating to food waste in the home³⁰. These questions can be used to inform the questionnaire design but may need to be tailored to ensure they meet the needs of an evaluation.

It is especially important to measure intermediate outcomes when the amount of HHFW in the home is not being measured (see sections 4.3.3 and 4.3.4). In such cases, it is important to see whether any of the metrics relating to intermediate outcomes have changed over the course of the intervention. Without this information, it is not possible to demonstrate any change in the real world because of the intervention. This makes it impossible to state how effective the intervention was.

4.3.3 Measuring levels of HHFW (final outcome)

Ideally all evaluations would include measurement of the amount of HHFW during the course of the intervention. This section provides guidance on this element of evaluation.

As discussed in section 3.2, it is important that the food waste that is measured is consistent with the aims of the intervention. Many interventions will focus on just the wasted food (edible parts), i.e. excluding inedible parts associated with food. In such cases, measurement could focus just on this fraction. Other interventions may only focus on a single food group (e.g. meat) or a single product (e.g. for a trial of a new packaging solution).

The decision on what fraction to measure is not always clear-cut. Measuring a smaller group of food requires more sorting (if waste compositional analysis

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC420179/>

²⁹ Guidance on cognitive testing can be found at:

<https://www2.gov.scot/Resource/Doc/175356/0091403.pdf>

³⁰ e.g. van Geffen, van Herpen, and van Trijp, 2017, "Quantified consumer insights on food waste: Pan-European research for quantified consumer food waste understanding" (see section 9.1): <https://eu-refresh.org/quantified-consumer-insights-food-waste>

is used), and there is usually a higher level of variability between households in the amount wasted. On the other hand, if the change in HHFW is concentrated in that category, the percentage reduction can be much higher, which can make it easier to detect the change. Statistical modelling can help in deciding the most appropriate course of action.

Whatever the decision about what HHFW is going to be measured, it should be clearly communicated. This should cover the following factors:

- Wasted food and / or inedible parts
- Destinations included (e.g. residual (general) waste, collections targeting food waste, sewer, fed to animals, home composting)
- Which food and drink categories are included
- Details on timescales (when the measurements are undertaken and how long the measurement lasts each time – see discussion on longevity in section 4.3.7)
- Geographical details (where are the sample of households located?)
- Other details relating to the sample: how they were selected if not drawn on a geographical basis (e.g. they were targeted through a workplace or a community group)

There are a range of methods available for measuring HHFW. However, for many of these methods their accuracy in monitoring change over the course of an intervention is doubtful. This includes survey-based methods (i.e. asking people to recall the amount of HHFW they produce), paper diaries, collecting HHFW in caddies and photo diaries. Most of these methods appear to substantially underestimate the amount of HHFW generated from households. Diaries have been shown to underestimate the amount of food by between 30% and 40%³¹; food caddies and photo diaries have similar levels of underestimation and surveys an even greater degree, as found in a recent REFRESH report³².

It is not yet known if the degree of underestimation for these methods varies over the course of an intervention. It is likely that interventions designed to engage people or otherwise increase the importance of food waste will influence the level of underestimation. Therefore, if these methods are used for evaluation purposes it will not be clear if any change measured i) reflects a real change in the amount of HHFW or ii) is an artefact of the measurement

31 E.g. WRAP (2013), [Methods used for Household Food and Drink Waste in the UK 2012](#)

32 Van Herpen, E., van der Lans, I., Nijenhuis-de Vries, M., Holthuysen, N., Kreme, S., 976 2016. Best practice measurement of household level food waste. EU Horizon 2020 REFRESH. <https://eu-refresh.org/best-practice-assessment-consumer-level-food-waste>

method (i.e. an interaction between the measurement method and the intervention itself).

It is possible that an intervention that increases the awareness of food waste in the home could lead to people recording more of their food waste (i.e. underestimating less) in post-intervention measurements, leading to an erroneous result that the intervention increases food waste. Until further research is undertaken to explore the suitability of these measurement methods in the context of evaluation, they are not recommended for evaluations of this nature.

However, some interventions do not engage people or raise the importance of food waste, e.g. increasing the shelf life of food products. In such cases, it may be possible to monitor using diaries, food caddies or photo-based methods. Efforts should be made to minimise bias relating from these methods, and results should be interpreted with caution (i.e. alongside other evaluation findings) until further research on their suitability is undertaken.

The remaining method for assessing household food waste is waste compositional analysis. This involves collecting waste streams that contain food waste from households, sorting through them to separate food from non-food material and weighing the amount of food. Some waste compositional studies further sub-classify the food waste to obtain detailed information about the types of food being wasted.

There is detailed guidance on the use of waste compositional analysis for monitoring interventions³³. This includes information on different types of compositional analysis, sample selection (including sample sizes required for different types of study), extrapolation and reporting.

The types of waste compositional analysis available are, in ascending order of cost:

- **Bulk Sampling:** typically occurs at a transfer station or disposal or recycling site; involves analysing waste mixed from a waste truck from a range of households
- **Small Area-Based Sampling:** Similar to bulk sampling but targets a specific physical area (e.g. a street or neighbourhood). This target could be the focus of an intervention.
- **Individual Sampling:** Waste samples are collected and analysed from individual (known) households. Can be linked to survey results from the households. It is the only practical way of undertaking WCA if the sample

³³ How to Measure Food Waste: A Guide for Measuring Food Waste from Households in Canada, National Zero Waste Council, June 2018
http://www.nzwc.ca/focus/food/Documents/LFHW_GuideToMeasuringFoodLossAndWaste.pdf

is not geographically defined (i.e. the evaluation needs to sample the waste only of workshop attendees).

More details on the advantages and drawbacks of each type of waste compositional analysis are provided in the previously mentioned National Zero Waste Council's guide.

Despite the accuracy of waste compositional analysis, it does have some potential draw-backs. It is often perceived to be more expensive than other methods, although different options are available: bulk sampling and small area-based sampling can be relatively low in cost. Because the amount of food generated by a single household varies over time, relatively large samples may be required to observe the effects of some interventions (especially those with more modest reductions in HHFW) – although this is an issue for all methods. It also does not measure HHFW going to certain discard routes, most notably: food and drink going down the sewer, being home composted or fed to animals in the home.

However, there are several studies that have successfully used waste compositional analysis to evaluate a HHFW-prevention intervention, including:

- Randomised control trial of an intervention in Canada using waste compositional analysis³⁴
- Community-focused initiatives in Worcester, UK, monitored using waste compositional analysis³⁵
- Campaign in West London (UK), evaluated using a questionnaire survey and waste compositional analysis³⁶

Furthermore, waste compositional analysis has underpinned long-term monitoring of levels of household food waste in the UK³⁷.

Before deciding on whether to measure food waste using waste compositional analysis, an estimate of the sample size required should be undertaken. This ensures that there is a sufficient number of households in the sample to make the comparisons required for the evaluation (e.g. comparing a sample before

³⁴ Van Der Werf, P. (2018), Developing and Testing a Novel Intervention to Reduce Household Food Waste., Ph.D. Thesis, University of Western Ontario: <https://ir.lib.uwo.ca/etd/5896/>

³⁵ Worcestershire County Council (2011), Reducing food waste through community focussed initiatives: http://www.wrap.org.uk/sites/files/wrap/2011.11_Worcestershire_CC_LFWW_2011_case_study.3e14035c.11397.pdf

³⁶ WRAP (2013), West London Food Waste Prevention Campaign Evaluation Report, http://www.wrap.org.uk/sites/files/wrap/West%20London%20Food%20Waste%20Campaign%20Evaluation%20Report_1.pdf

³⁷ WRAP (2016), Synthesis of Food Waste Compositional Data 2014 & 2015, http://www.wrap.org.uk/sites/files/wrap/Synthesis_of_Food_Waste_2014-2015.pdf

and after the intervention). If the sample is too small, it means that any effect of the intervention cannot be detected amongst the 'noise' in the data.

To determine an appropriate sample size, the following information is required:

- The mean amount of food waste relating to the fraction in question (e.g. all food waste, only bakery, one product)
- The distribution and variation of food waste (ideally between households and over time for individual households)
- The anticipated reduction by the intervention on the fraction of food waste being investigated. At the design stage, this can be an educated guess.

Information for the first two bullets can be derived from existing datasets on household food waste. Someone with statistical knowledge should be consulted to help determine the minimum sample size to detect the anticipated reduction in food waste, via the use of power calculations or statistical modelling.

It is important to remember that people may drop out of an evaluation during its course (attrition), thereby reducing the final sample size. This can be counteracted by starting the evaluation with more people (or households) than are required for the analysis. The level of attrition will be dependent on a range of factors – e.g. the types of measurement method used, the nature of the people and households in the sample, the intervention itself and any incentives offered to research participants. Drawing on similar research and discussion with experts is recommended to estimate the level of attrition that is likely in your evaluation.

Further guidance on all aspects of measuring food waste – including descriptions of the different methods – is provided in the *Food Loss and Waste Accounting and Reporting Standard*³⁸.

In summary, studies evaluating food-waste-prevention initiatives in the home should aspire to measure the change in food waste in a sample of homes over the course of the intervention (and after it has completed to investigate longevity effects – see section 4.3.7). This should be undertaken using an objective measure, such as waste compositional analysis. However, this may not always be practically possible, so alternatives are presented in the following section.

³⁸ <http://flwprotocol.org/>

4.3.4 Alternatives to measuring HHFW using waste compositional analysis

There are limited alternatives to measuring the change in HHFW via waste compositional analysis. The following, although not ideal, may be sufficiently low cost and may be better than not estimating the change of HHFW at all:

Measure up to intermediate outcomes and then effect on HHFW: this involves measuring the intermediate outcomes of the intervention (e.g. change in behaviour, change in shelf life of product) and then modelling the effect of these changes on household food waste. Methods for this modelling are being developed: for example, The Milk Model³⁹ study demonstrated that using simulation modelling for this purpose looked promising. At the time of writing, this work is being further developed by WRAP and Sheffield University to include a wider range of products and household characteristics; this will be ready for wider use towards the end of 2019. Any modelling or use of assumptions should be clearly stated in the evaluation report.

Use diaries, photos or caddies to estimate HHFW: as discussed in section 4.3.3, this should only be undertaken if there is unlikely to be a change in how people respond to these measurement methods (i.e. the degree of underestimation is likely to be stable). Depending on the research design, there is also the possibility that the measurement method could lead to HHFW prevention: e.g. filling in a diary may cause people to reflect on their HHFW and take steps to prevent it.

For these types of research, consideration should be given to the following:

- Minimising selection bias using a robust sample selection procedure
- Maximising sample retention, via incentives and an easy-to-use 'measurement instrument' (e.g. diary)
- High levels of accuracy in participants' recording of food waste, using a well-designed measurement instrument

Useful references for those considering these methods include a recent REFRESH report covering surveys, diaries, caddies and photos⁴⁰, a recent

³⁹ WRAP (2013), The Milk Model: Simulating Food Waste in the Home, <http://www.wrap.org.uk/sites/files/wrap/Milk%20Model%20report.pdf>

⁴⁰ Van Herpen, E., van der Lans, I., Nijenhuis-de Vries, M., Holthuysen, N., Kreme, S., 976 2016. Best practice measurement of household level food waste. EU Horizon 2020 REFRESH. <https://eu-refresh.org/best-practice-assessment-consumer-level-food-waste>

paper on photographic methods for quantifying plate waste⁴¹ and an example of a diary designed for quantification purposes⁴².

4.3.5 Quantifying impacts

Many evaluations would like to make statements about the environmental, economic and / or social impacts of preventing HHFW.

For environmental and economic impacts, it is not usually possible to measure these directly. For example, it is hard to directly measure the reduction in greenhouse gas emissions that ensues from HHFW prevention. Most estimates of environmental and economic impacts rely on 'factors' being applied to the weight of HHFW prevented. Environmental factors are usually obtained from Life Cycle Assessments (LCAs), while economic factors are often derived from the retail price of food. Factors are often specific to a type of food and represent the amount of, for example, greenhouse gas emissions associated with a kilogramme of, say, bread⁴³.

This type of approach is relatively straight-forward to apply. However, it can overlook complexities, such as the rebound effect⁴⁴ or interactions between different parts of the supply chain⁴⁵. Currently the impact of these complexities is uncertain, and therefore the simpler approach can be used for obtaining approximate estimates. However, future research may provide more sophisticated tools for assessing the impact of HHFW.

For social impacts – e.g. on people's access to food – there are no studies using the 'factor' approach described above. If the wider social impacts of a HHFW prevention intervention are important to the study, then bespoke evaluation will be required to understand these interactions, mostly likely requiring qualitative research and self-reported information on social impacts from questionnaires.

⁴¹ Roe B.E., J.W. Apolzan, D. Qi, H.R. Allen, and C.K. Martin. 2018. Plate waste of adults in the United States measured in free-living conditions. PLoS ONE. 13(2).

⁴² <http://www.wrap.org.uk/sites/files/wrap/Kitchen-Diary-2012-Final-Low-Res.pdf>

⁴³ Examples of studies containing this type of approach include: WRAP (2011), *The water and carbon footprint of household food and drink waste in the UK*

http://www.wrap.org.uk/sites/files/wrap/Water_and_Carbon_Footprint_report_Final.pdf

WRAP (2013), *Methods used for Household Food and Drink Waste in the UK 2012* (Chapter 9)

<http://www.wrap.org.uk/sites/files/wrap/Methods%20Annex%20Report%20v2.pdf>

⁴⁴ Salemdeeb R., et al. (2017), *A holistic approach to the environmental evaluation of food waste prevention*, *Waste Management*, 59, 442-450, <https://www.sciencedirect.com/science/article/pii/S0956053X16305463>

⁴⁵ Campoy-Muñoz, P., M.A. Cardenete, and M.C. Delgado. 2017. "Economic impact assessment of food waste reduction on European countries through social accounting matrices." *Resources, Conservation and Recycling* 122:202–209.

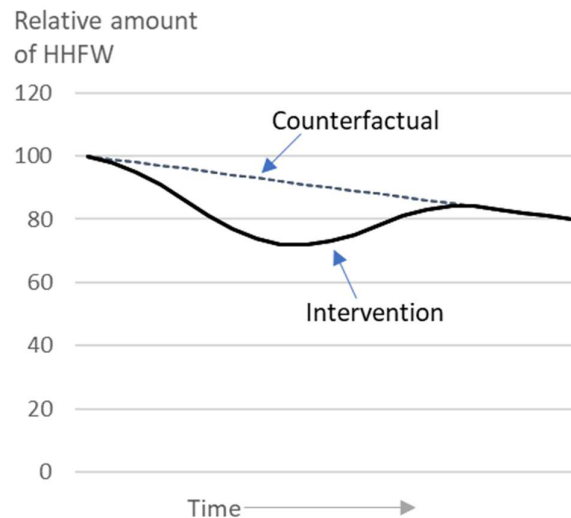
An in-depth discussion of these topics can be found in a recent technical report from the Commission for Environmental Cooperation⁴⁶.

For information on the impact of interventions preventing HHFW, methods and assumptions used to calculate the environmental, social and economic impacts should be clearly reported.

4.3.6 Understanding the counterfactual

An impact evaluation attempts to assess the outcomes and impact of an intervention. In effect, an evaluation attempts to compare – for the same household and time period – the situation in which the intervention has been deployed to one where the evaluation is absent (the 'counterfactual'). In the hypothetical example shown in Figure 2, the counterfactual shows a decreasing level of food waste over time. This could be for a number of reasons – e.g. rising food prices. The effect of the intervention is the difference between the counterfactual and the intervention, which varies over time.

Figure 2: Illustration of levels of household food waste for an intervention and counterfactual



The counterfactual can be determined in a number of ways, listed below in decreasing order of robustness⁴⁷:

⁴⁶ See chapter 4 of: <http://www3.cec.org/islandora/en/item/11813-technical-report-quantifying-food-loss-and-waste-and-its-impacts>

⁴⁷ Adapted from the Maryland Scientific Scale: <https://whatworksgrowth.org/resources/the-scientific-maryland-scale/>

- **Randomised control trial (RCT):** in which some households are exposed to the intervention and others are not (the control group). Each household involved in the study should be randomly allocated to either the intervention or control group⁴⁸.

There are many variants on RCTs including the staggered-start: e.g. many people apply to take part in a waste-prevention workshop, half of this group (chosen randomly) undertake the intervention in the first wave. During this time, the second half of the group don't undertake the workshop but act as a control. This second half then undertake the workshop at a later date.

RCTs are often performed with measurement of HHFW (and other variables) before and after the intervention (i.e. a pre and post design). An alternative approach where households are randomly assigned into intervention and control groups, it is possible to determine the impact of the intervention by only measuring HHFW after the intervention, and then comparing the levels for the two groups.

- **Quasi-randomised trial:** similar to an RCT, some households will be exposed to the intervention and others are not in a way that is caused by external factors (i.e. those not deliberately imposed by those running the intervention or the evaluation). Although these may not be truly random, these factors should be sufficiently close to random for this type of trial. Additional analysis is required to demonstrate that the external factors (i.e. those influencing whether a household is exposed to the intervention) are unlikely to influence the results. An example of this is where an intervention is rolled out sequentially in different regions of a country; as each region begins the intervention, the regions still without the intervention could act as a control. It would need to be demonstrated that the order of the regions receiving the intervention did not influence the results.
- **Comparison between intervention group and non-random comparison group:** this type of counterfactual involves a comparison group that has not been randomly selected but chosen to be similar to the intervention group. For example, if a supermarket is running an intervention for its customers (e.g. in store), then a comparison group could be selected from customers to a similar store that is not running the intervention (e.g. a matched comparison store). The comparison group would need to be 'matched' to the intervention group on variables important to household food waste: for example, age, structure of the household (e.g. presence of children), time devoted to food management and cooking, etc.

In some situations, it is possible to model the counterfactual using known relationships between levels of household food waste and other factors. An

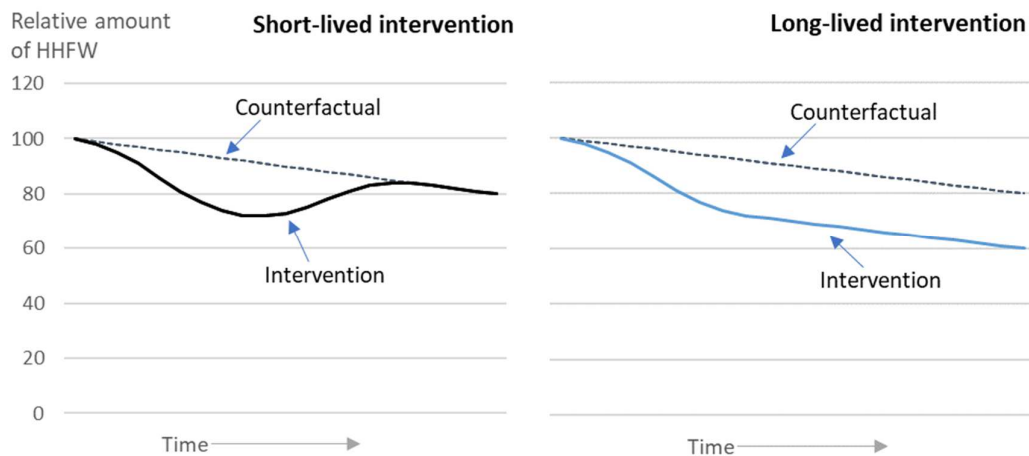
⁴⁸ Expert advice should be sought on methods to achieve this

example is the econometric modelling performed by WRAP to investigate the influence of food prices and income levels on levels of household food waste in the UK⁴⁹. However, this type of approach is approximate, often with much uncertainty around the estimates of the counterfactual. More research is required before modelling counterfactuals is a viable option for evaluating HHFW prevention interventions.

4.3.7 Longevity of outcomes and impacts

To understand the full impact of an intervention, it is important to know how long any changes continue. If, at the end of an intervention, there was a reduction in HHFW of 20%, there is a much bigger impact if this persisted for many years (see right-hand diagram in Figure 3), compared to a case where the level of food waste reverted to its original level within a matter of months (see left-hand side of Figure 3).

Figure 3: Examples of short-lived (left) and long-lived (right) interventions



To date, very few studies have investigated the 'longevity' of HHFW interventions (i.e. their lasting effect). Therefore, it is not known how different types of intervention compare when looking over their full lifetime.

Therefore, when designing evaluations, it is important to continue monitoring after the intervention has stopped (e.g. after the campaign has stopped running, after the initial introduction of new packaging). The exact length of time to assess longevity will depend on the intervention and may need to respond to emerging data (e.g. if there is still a strong effect of the

⁴⁹ WRAP (2013), *Econometric modelling and household food waste*
<http://www.wrap.org.uk/content/econometric-modelling-and-household-food-waste>

intervention after 3 months, it might be useful to re-measure after 12 months).

Barriers to measuring longevity include additional cost and resources, and a potential delay in the full findings emerging. Given the importance of this information, all involved in the intervention and evaluation should seek additional funds to support longevity measurements. To counteract the potential delay before the full findings are published, staged release of results could be undertaken.

4.4 Data protection

The evaluation will be collecting information relating to people and households. Therefore, it is important that the evaluation complies with relevant laws on data protection and data security. Relevant to projects involving citizens of the European Union is the General Data Protection Regulation (GDPR)⁵⁰.

Key considerations for this type of evaluation include:

- Sharing data within multi-partner research teams
- Sharing data with third parties (e.g. survey contractors)
- Re-contacting intervention participants (e.g. those attending a workshop) to obtain feedback or participate in evaluation-based research

Steps that are especially important for this type of evaluation include getting informed and explicit consent from research participants prior to any research, informing people whose data is being collected how it will be used and protected, and setting up and documenting processes to demonstrate that these have been done. Specialists in data protections (e.g. the organisation's Data Protection Officer) should be consulted by the evaluation team to ensure that the regulations are being followed.

Given this, data protection should be considered **before** the evaluation starts. Failure to do so could lead to information crucial to the evaluation not being collected because the correct data-protection steps have not been followed. This can prevent the evaluation from being able to answer its evaluation questions and therefore it will be of limited value to its audience.

4.5 Outputs of the evaluation-development process

By the end of the process of developing the evaluation, the following are usually required:

⁵⁰ <https://eugdpr.org/>

- A clear understanding of the intervention, often incorporating a logic map or a similar tool
- Evaluation questions – a set of priority questions that the evaluation seeks to answer.
- An evaluation plan that includes what information – qualitative and quantitative – will be collected, how it will be collected. It includes who will be conducting each piece of research and timings. For each piece of information, there should be a clear link to the evaluation questions.
- The evaluation plan also includes how the findings will be used, published and communicated.

The elements listed above should be consistent with one another. For example, the evaluation plan should be able to answer the evaluation questions and be appropriate given the nature and context of the evaluation. Inconsistencies should be made to resolved before the evaluation is undertaken.

As mentioned in the next chapter, most evaluations need to be flexible. Nevertheless, an evaluation has a greater chance of producing useful information if it has been well planned and there is a clear link between the different elements of the evaluation.

5 Implement and disseminate the evaluation findings

This section provides summary advice relating to the implementation and dissemination of an evaluation's findings. Further guidance can be found in evaluation references, e.g. Chapter 10 of the Magenta Book⁵¹.

5.1 Implementing the evaluation

Implementing the evaluation primarily involves implementing the evaluation plan: collecting, analysing and reporting the information that forms the evaluation.

However, implementation usually requires flexibility. The intervention may not be implemented exactly how it was envisaged during the development stage: budgets may change, delivery partners may change, lessons from early stages of the evaluation may improve the later stages of implementation. In addition, the understanding of the intervention (its theory) may also evolve as the evaluation progresses.

Therefore, the evaluation also requires flexibility: the evaluation will need to adapt to answer the evaluation questions. It is good practice to report both the initial evaluation plan, but also the changes that occurred and why these changes were made.

5.2 Dissemination of findings

To allow the evaluation to be used by a range of audiences, it is useful for the following to be disseminated:

- **Types and amounts of resources** and other inputs used during the intervention. In addition, it is useful to provide an indication of how these resources could change in the future: e.g. the interventions could be deployed at lower cost in the future, given that the materials have been developed.
- **Description of the intervention:** including what was done, when, and who was involved. In addition, the information relevant to the context of the intervention should also be reported, as this can influence the outcomes of the intervention⁵². Reproductions of the 'materials' developed

⁵¹ HM Treasury 2011, The Magenta Book, Guidance for evaluation: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

⁵² This could include details of how the intervention participants compare to the general population, any significant background or events that were occurring during the intervention / evaluation (e.g. economic conditions such as recession, changes in food prices, trends in food consumption or culture)

for the intervention (e.g. campaign leaflets, changes in packaging) should be included.

- **Evaluation methodology:** how evaluation information was gathered and analysed, including details of sampling approaches and characteristics of achieved samples. There should be sufficient detail that it could be reproduced by others. Ideally, the research instruments (e.g. survey questions, diary used) are made available so that studies can be replicated, and results compared.
- **Results:** All quantitative and qualitative information relating to the outputs, outcomes (intermediate and final) and impacts (if calculated). This should be presented clearly and include, where appropriate, comparison with control groups or other measures of the counterfactual.
- **Discussion and recommendations:** an objective assessment of the elements of the intervention that worked well and for whom (led to intended outcomes and impacts), and those that didn't. Recommendations can cover whether the intervention should be stopped, modified or rolled out more widely. Recommendations can also include whether the intervention is likely to succeed in other setting, for example, in other countries. The limitations of the evaluation should be acknowledged alongside how these affect the results.
- **Funding for the evaluation and any conflicts of interest** allowing the reader to assess whether these have the potential to influence the findings.

6 Good-practice tips

The following reiterates some of the key points of good practice for an evaluation of household food waste.

Start evaluations early: if evaluation is considered early in the development of an intervention, the more likely the evaluation is to proceed, be of an appropriate design, and provide useful information for those implementing the intervention or those considering a similar intervention. In general, development of the evaluation should mirror the development of the intervention. For example, the budget for the evaluation should be considered alongside the budget for the intervention. The evaluation should be developed as the intervention is being designed.

Collaborate: robust evaluations should be developed and delivered with input from a range of people, usually including those delivering the intervention, evaluation specialists and researchers focusing on food waste.

Understand the intervention using logic maps and sharing of information between the evaluators and those deploying the intervention. This should include discussions on who the target audience is and how the intervention is designed for the context in which it is going to be deployed (Chapter 3).

Develop an evaluation plan: this should include a logic map of the intervention, the evaluation questions, a plan of what information will be gathered, when and by whom. It includes how the information will be analysed, documented, shared and communicated (Chapter 4).

Use a mix of information and methods to inform the evaluation: the best evaluations are usually informed by a mix of quantitative and qualitative information from across the logic map. Numerical data is supplemented by information gathered from observation and interviews (section 4.3).

Try to measure HHFW accurately. If at all possible, use waste compositional analysis, as it is the measurement method least subject to bias. In addition, ensure the sample sizes is sufficiently large to be able to detect the changes you're looking for, remembering that a proportion of people will drop out of the evaluation over its duration (section 4.3.3).

Include a counterfactual: Where possible compare households exposed to the intervention with a carefully chosen comparison group (see 4.3.6).

Consider long-term change: many evaluations consider the outcomes and impacts during and immediately after the intervention. However, to understand the full impact of the intervention, it is necessary to know if these impacts persist after the intervention or decrease over time (section 4.3.7).

Publish and disseminate findings including the types and amounts of resources used in the intervention, description of the intervention, evaluation methodology, results, discussion and recommendations. In addition, it should include where the funding for the evaluation came from and any possible conflicts of interest (section 5.2).

7 Conclusion

There is a clear need to understand the effectiveness of different policies and interventions designed to reduce the amount of household food waste (HHFW). Currently, a lack of evidence hinders this understanding – in particular, a lack of comparable evaluation studies that robustly measure the impact on HHFW of the intervention on food waste.

To tackle this lack of evidence, the current guidance has been developed to help stimulate future good-quality evaluation studies. This guidance covers three main stages:

- Understanding the intervention (Chapter 3)
- Developing the evaluation plan (Chapter 4)
- Implementing the evaluation plan and disseminating the findings (Chapter 5)

The guidance draws on a number of evaluation references, that provide an overarching framework for evaluation and guidance on each step. However, evaluating HHFW interventions has a number of specific difficulties to overcome:

- The underlying phenomenon is complex, involve many interacting activities in the home
- Measurement of HHFW is not straight-forward: many common methods (surveys and diaries) are prone to considerable biases, making them unsuitable for most evaluation purposes
- Sample sizes required to observe effects can be relatively high due to high variability in amounts of HHFW generated
- Access to intervention participants, especially if they have to be contacted via an intermediary organisation and / or aren't within a bounded geographical area.
- The effectiveness of interventions is context-specific: what prevents HHFW in one household may not work in another; what prevents HHFW in one country may not work in another.

This document gives guidance on how to overcome these issues and highlights a range of references and resources for these specific issues.

In the future, use of this guidance should lead to a step-change in the quality of studies evaluating HHFW. This should provide evidence for policy makers and other decision makers to select the most appropriate approaches, so that they are able to reduce the amount of food wasted from households in a cost-effective manner.

8 Further reading

General advice on evaluation can be found on the following websites:

<https://betterevaluation.org/>

<https://www.evaluation.org.uk/index.php/news-resources/ukes-publications/46-ukes-guidelines-for-good-practice-in-evaluation>

Most Member States will have general guidance on evaluating government policy. For example:

HM Treasury 2011, *The Magenta Book, Guidance for evaluation*:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

Some types of intervention have more specific guidance. For example, communications:

Evaluation Framework 2.0, Government Communication Service, June 2018: <https://gcs.civilservice.gov.uk/guidance/evaluation/tools-and-resources/>

... and food waste across the supply chain:

Burgos et al. (2016) Policy Evaluation Framework, FUSIONS report: <http://www.eu-fusions.org/index.php/download?download=255:policy-evaluation-framework>

Caldeira, C., De Laurentiis, V., Sala, S., Assessment of food waste prevention actions: development of an evaluation framework to assess performance of food waste prevention actions, JRC Technical report. Interim document - Limited distribution, 2019, JRC115987 (due to be published later in 2019)

Guidance for logic mapping includes *The Magenta Book* (see above) and:

Tavistock Institute (for the Department of Transport), 2010 *Logic Mapping: Hints and Tips Guide*:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/3817/logicmapping.pdf

Guidance on measuring food waste in the home via waste compositional analysis can be found in this report:

How to Measure Food Waste: *A Guide for Measuring Food Waste from Households in Canada*, National Zero Waste Council, June 2018

http://www.nzwc.ca/focus/food/Documents/LFHW_GuideToMeasuringFoodLossesAndWaste.pdf

Furthermore, studies looking at alternative measurement methods include:

Van Herpen et al., 976 2016. *Best practice measurement of household level food waste*. EU Horizon 2020 REFRESH. <https://eu-refresh.org/best-practice-assessment-consumer-level-food-waste>

Roe et al. 2018. *Plate waste of adults in the United States measured in free-living conditions*. PLoS ONE. 13(2).

These two papers outline the relative lack of evaluation studies relating to prevention of HHFW in the home, and the lack of standardization (that this guidance hopes to address):

Reynolds et al. 2019, Consumption-stage food waste reduction interventions – what works and how to design better interventions, *Food Policy*, 83, 7-27

Stöckli, Niklaus, Dorn, 2018: *Call for testing interventions to prevent consumer food waste*, *Resources, Conservation and Recycling*, 136, 445-462

This issue is further discussed, alongside a summary of the state of knowledge relating to HHFW in this policy briefing and the accompanying document providing further background:

Wunder, 2019. *REFRESH Policy Brief: Reducing consumer food waste*. EU Horizon 2020 REFRESH. <https://eu-refresh.org/node/908/>

Wunder et al. (2019). *Policies against consumer food waste. Policy options for behaviour change including public campaigns*. EU Horizon 2020 REFRESH <https://eu-refresh.org/policies-against-consumer-food-waste>