

# Step 2: Plan

## Introductory reading

### Introduction

The following document provides a short introduction to Step 2 of TASO's Monitoring and Evaluation Framework (MEF). This step is called 'Planning' and will cover reading for the following topics:

1. Identifying Research Questions
2. Identifying Outcome Measures
3. Selecting a Research Method
4. Developing an Analysis Strategy
5. Creating a Research Protocol
6. Self-Assessing Evaluation Security

### 1. Identifying Research Questions

In Step 2 of the evaluation process, you will need to use your Theory of Change to develop the questions that your evaluation will seek to answer. These overarching questions will determine the scope and approach of your evaluation.

The first research question should be about the causal impact of the intervention or scheme:

Did [scheme] increase [outcome] among [group]?

*For example:*

- Did the residential summer school increase acceptances to highly selective universities among high-achieving students from low-income backgrounds?
- Did Welcome Week improve first year attainment among first-year undergraduates?

You might also wish to have secondary research questions focusing on the impact for specific groups, or for intermediate outcomes:

- Did Welcome Week improve first year attainment among first year undergraduates from widening participation backgrounds?
- Did Welcome Week increase a sense of belonging among first-year undergraduates?

You may also wish to have research questions relating to other effects of the intervention, or about the way it was implemented and experienced by recipients (mapping to the process elements of your Theory of Change – steps 5-7) such as:

- Was the initiative delivered the way we expected?
- Are we targeting the right students?
- What was the cost-effectiveness of the initiative?

To help formulate your questions, you should also consider:

- Who will use the findings and how?
- What do stakeholders need to learn from the evaluation?
- What questions will you be able to answer and when?

## 2. Identifying Outcome Measures

Once you have established your research questions, you will need to consider which outcome measures best enable you to answer them and demonstrate success. The measures should link closely with the process, outcomes and impact you have recorded in your Theory of Change. A simple way to think about which measures to select is:

“I’ll know [outcome reached] when I see [indicator]”

The **Common Outcome Measures table** sets out common outcome indicators for initiatives at each stage in the student life-cycle, from Key Stage 3 through to post-graduation. The framework supports the identification of outcome measures or indicators, relating to specific objectives, to support the measurement of progress or achievement of outcomes.

If it is necessary to develop new indicators outside those in your Common Outcome Measure table, it’s worth considering the below hierarchy of measures based on their reliability and validity:

0. Output only,
1. Self-report subjective (e.g. perceived knowledge),
2. Self-report objective (e.g. actual knowledge),
3. Validated scales (e.g. from academic research, externally-administered tests),
4. Interim or proxy outcome (e.g. GCSE selections, sign-ups to events), or
5. Core impact (e.g. A level attainment, university acceptances, continuation).

Generally, we should aim to be focusing evaluations on measures at the higher end of this scale (i.e. 3 and above).

### 3. Selecting a Research Method

#### *Impact evaluation*

There are many different methods that can be used to try and understand both whether your initiative is having an impact, and how it's operating in practice. In this section, we focus on the primary research method – that is, the research method being used to investigate your primary research question, which will enable you to measure the causal impact of your initiative on an outcome.

Overall, some research methods are better suited to this question than others. Following the **OfS' Standards of Evidence**, we conceptualise three levels of impact evaluation:

- Monitoring,
- Comparing and;
- Identifying.

Over time, we would encourage all programmes across the sector to move towards having Level 2 or, where feasible, Level 3 impact evaluations. However, this process may occur over a number of years, especially for new or complex initiatives.

The diagram below summarises some of the key research methods at each level. You can also **download the research methods at each level**.

#### ***Level 1 – Monitor***

***We have a coherent strategy and activities are selected to contribute to that strategy. We know why we expect particular activities to work (based on a Theory of Change and secondary research) and we are tracking participants' outcomes and experiences.***

Level 1 evaluation is a basic expectation of all services and schemes. Initiative owners should lead on planning for monitoring the outcomes of participants and service users. This includes secondary research to guide initiative development, tracking destinations of participants, and conducting post-initiative research to gauge the position of participants post-service. Institutional evaluation teams, if available, can advise on the development of a Level 1 evaluation approach and may in some cases be able to support delivery (for example, conducting focus groups or data analysis), but monitoring is the responsibility of initiative owners.

#### ***Level 2 – Compare***

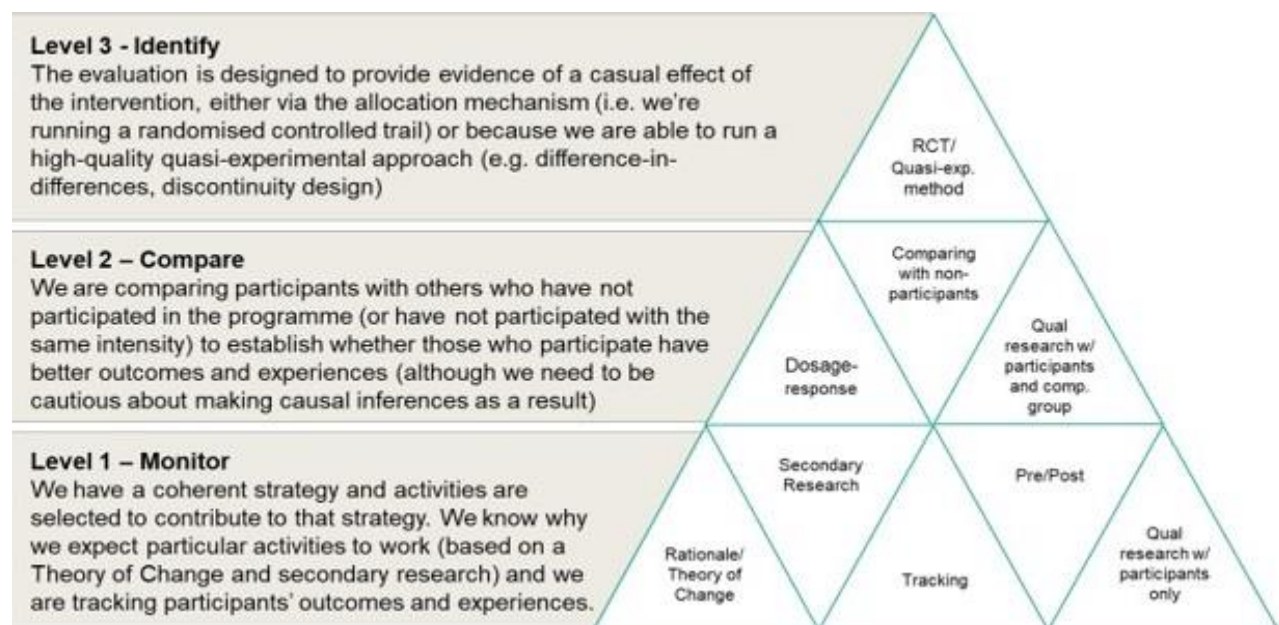
***We are comparing participants with others who have not participated in the programme to establish whether those who participate have better outcomes and experiences.***

Level 2 evaluation should, over time, be feasible for all services and schemes. At this level, institutional evaluation teams, if available, will lead on evaluation, agreeing a research approach with the initiative owners and co-drafting the Research Protocol.

### **Level 3 – Identify**

**The evaluation is designed to provide evidence of a causal effect of the intervention, either via the allocation mechanism or because we are able to run a high-quality quasi-experimental approach.**

Level 3 evaluation is the goal for some services and schemes; however, the form of the evaluation, the timelines, and whether it is ultimately feasible will vary. It is important to note that in some cases a high-quality Level 2 evaluation will provide better evidence of an impact than the available or feasible Level 3 evaluation approaches. Institutional evaluation teams, if available, should always lead on Level 3 evaluations, agreeing a research approach, drafting the Research Protocol and conducting the evaluation. Initiative owners will be closely involved at all stages to ensure that the evaluation design doesn't impact on service delivery.



The diagram above summarises some of the key research methods at each level.

### **Process evaluation**

**At this stage you should also consider the best way to collect data about the way the initiative worked, whether everything went to plan, and how it felt to participants and partners. The methodology guidelines also contain an overview of common process evaluation methods.**

## 4. Developing an Analysis Strategy

Based on your research method, you should consider how the data is going to be analysed. This is equally important – if not more so – for qualitative and process evaluations, where there are likely to be more research questions with a less direct link between the question, method and analysis strategy. For instance, if you are conducting focus groups or interviews, will you take notes or will they be recorded and transcribed? In the latter case, how will you convert them? There are a range of different methodologies and software that can be used, and conducting robust qualitative research is as difficult as conducting robust quantitative research.

Deciding your analysis strategy in advance reduces the temptation to cut the findings in a way that supports what you would like to find. It also gives you a roadmap through the data, which can sometimes be overwhelming in the number of options it presents, and help maintain focus on the key questions you wanted to answer.

For Level 2 and 3 evaluations, the institutional evaluation team, if available, will work with you on the best analysis strategy, and will provide advice and guidance on Level 1, survey and qualitative research.

## 5. Creating a Research Protocol

A Research Protocol is a written document that describes the overall approach that will be used throughout your intervention, including its evaluation.

*A Research Protocol is important because it:*

- Lays out a cohesive approach to your planning, implementation and evaluation
- Documents your processes and helps create a shared understanding of aims and results
- Helps anticipate and mitigate potential challenges
- Forms a basis for the management of the project and the assessment of its overall success
- Documents the practicalities of implementation

*Reasons for creating a Research Protocol include:*

1. Setting out what you are going to do in advance is opportunity to flush out any challenges and barriers before going into the field.
2. Writing a detailed protocol allows others to replicate your intervention and evaluation methodology, which is an important aspect of contributing to the broader research community.

3. Setting out your rationale and expectations for the research, and your analysis plan, before doing the research gives your results additional credibility.

The protocol should be written as if it's going to end up in the hands of someone who knows very little about your organisation, the reason for the research, or the intervention. This is to future-proof the protocol, but also to ensure that you document all your thinking and the decisions you have made along the way.

- **[Download our Research Protocol Template](#)**

## **6. Self-Assessing Evaluation Security**

Based on the decisions made around the evaluation, you will be able to assess the security of your evaluation – that is, how confident you can be when making claims about the findings. The most robust evaluations with large samples, low attrition levels and no threats to validity will receive the highest score of 5/5. However, it is worth bearing in mind that in many cases it will not be feasible to achieve a score this high, due to the nature of the research questions and the subsequent evaluation methods used to answer them. Your overall rating will be calculated by taking the average score of each section in the table below.



	0 (Low)	1	2	3	4	5 (Excellent)	Rating
Research Design	No comparison group (Level 1)	Comparison group, poor/ no matching (Level 2)	Matched comparison (Level 2)	Well-matched comparison (Level 2)	Quasi-experimental approach (Level 3)	Randomised Controlled Trial (Level 3)	
	See Section 4.3 and Annexe 2 for more detail about research designs.						
Sample (per arm)	< 50	50 - 100	100 - 500	500 - 1000	1000 – 2000	> 2000	
Outcome measure	Output only	Self-report subjective	Self-report objective	Validated scales	Interim or proxy outcome	Core impact measure	
	See Section 4.2 and Annexe 1 for more detail on the common core outcomes.						
Attrition	≥50%	<50%	<40%	<30%	<20%	<10%	
	The proportion of the cases we’re expecting we’ll be unable to collect outcome data for, or who will withdraw from the intervention while it’s running.						
Validity	Major threats to validity	Some threats (e.g. major deviations from protocol)		Minor risks (e.g. small deviations from protocol)		No threats to validity	
	Any other threats to the validity of <u>the</u> : e.g. is outcome data collected inconsistently, might control individuals get some of the intervention, are the assumptions of the research method unjustified?						
OVERALL RATING	Sum individual ratings and divide by 5						/ 5