

Five essential building blocks for clearer findings

A practical guide to
analysis and reporting
that leads to action



Jim Barratt
October 2026

Foreword

Working as a freelance evaluator, I've had to develop and refine a way to make sense of evidence and turn it into reports that busy people can actually use.

I've completed more than 175 consultancy assignments since 2006, learning the craft of analysis and reporting along the way and honing my approach rather than following a single, ready-made template.

That experience has taught me that while every evaluation is different, the underlying challenge is often the same: how to move from a mass of evidence to a clear, defensible judgement that someone can act on.

A common concern I hear from people new to evaluation is that analysis and reporting can feel both daunting and risky.

I feel it too, even after all these years. But analysis and reporting are essential parts of the job, and also the most visible.

Whatever thinking went into the evaluation design and data collection methods, the commissioner or reader often only sees your effort recorded in the final report.

If the analysis feels unstructured or the report is hard to follow, it does not matter how good the underlying work was.

The credibility and usefulness of the evaluation is judged on the final output, whether that is a full written report, an infographic or a slide deck presentation of your findings.

There is no shortage of excellent guidance about planning an evaluation, building a Theory of Change, developing rubrics, choosing appropriate methods and data collection tools etc.

Harder to come by is practical help with managing your analysis and writing process so you can present a clear judgement without burying the reader in distracting detail or neglecting key parts of your story.

This guide is my attempt to make that crucial aspect of the craft more open to scrutiny, easier to implement and less of a chore overall.

My hope is that these building blocks help you not only organise your own thinking and writing, but also produce findings that commissioners, funders and other readers of your evaluation reporting can use with confidence.

I intend to refine and update the guide as my thinking develops, and I welcome your feedback.

Please get in touch if you have any suggestions, comments or questions.

Jim Barratt

jim@biggerpictureresearch.com

Contents

About this guide.....	3
Introducing the five building blocks.....	6
Scene Setting.....	13
Evidence.....	18
Assessment.....	23
Significance.....	27
Implications.....	31
Real world examples of the building blocks in action.....	35
Glossary.....	39
Further reading.....	40
Annex: Example project documentation.....	41

Acknowledgments

I want to thank Catherine O’Leary, Tudor Barratt and Tim Bidey, whose feedback on earlier drafts offered the right combination of encouragement, validation and helpful suggestions.

Any shortcomings in this guide are, of course, my responsibility alone.

Cover illustration by Arianne Barratt

*Five essential building blocks for clearer findings:
A practical guide to analysis and reporting that leads to action*
© 2026 by Jim Barratt
is licensed under Creative Commons Attribution-ShareAlike
4.0 International

To view a copy of this license, visit <https://creativecommons.org/licenses/by-sa/4.0/>



Third-party report names and references cited in this guide remain the property of their original authors.

AI declaration

I used generative AI to create the outlines of two fictional projects used as examples in this guide. I did this to avoid inadvertently sharing examples that resembled projects I worked on during 20 years of consulting.

I conducted an online search to check these AI generated examples were not copied from any identifiable real-world projects, and I am satisfied they are likely to be generic composites based on project details the AI model was trained on (although I cannot know this for sure).

As a further step, I also modified the project descriptions and example documentation to ensure they were properly fictionalised with human input, and fit for purpose in the context of this guide (as the original AI generated examples contained discrepancies and irregularities). Any resemblance to actual past or planned projects is wholly unintended.

About this guide

Evaluation analysis can suffer from the lack of a clear structure. Without one, it is easy for it to sprawl into everything that is merely interesting rather than what is relevant and important, or to get fixed on a judgement before the evidence has been thoroughly explored.

A related problem affects the write-up that follows. The desire to demonstrate methodological rigour often leads to reports built around the research process rather than what stakeholders need to know in order to act on the findings.

This can result in reports that bury key findings in technical detail, or that assume readers will understand why conclusions matter and what to do about them.

In response, this guide sets out five building blocks that can be used to structure your evaluative thinking and reasoning:



In simple terms, the blocks are ways of thinking about how to interpret and present your findings.

Used in combination, they provide a framework for weighing evidence, forming a judgement and communicating it effectively so it leads to action. If you apply the building blocks during analysis, it should make the reporting process much more straightforward.

There is no clever theory behind this common sense approach. Because it has developed out of my own consultancy work, it is inherently practical.

Throughout the guide, I use the term 'project' to refer to any initiative that is being evaluated. Where you see this term, you can also read it as 'programme', or any other wording that refers to a planned intervention to address a particular problem or challenge in the world.

You will notice that the five building blocks are not summarised with a neat acronym, and this is deliberate. The approach is designed to be modular, so the order in which you use the building blocks should follow on from the task in front of you, not a fixed formula.

An acronym would imply a single correct ordering, and no single sequence fits every circumstance. This is explained in more detail later on.

What 'analysis' means in this guide

Analysis is used here to mean the process of making sense of evaluation evidence and reaching a conclusion about it.

This guide is not concerned with the technical aspects of analysing data: for example, statistical analysis of quantitative data, or the coding and thematic analysis of qualitative data.

Rather, it is concerned with what happens once relevant evidence has been assembled: how it is considered against the evaluation questions, indicators and Theory of Change; how different sources are brought together; and how this leads to a credible assessment.

What 'reporting' means in this guide

Reporting is used here to mean the process of communicating your findings to others. This may take many forms, including a formal evaluation report, an executive summary, a slide deck, an infographic or a briefing.

The guide does not offer advice on writing style, presentation or graphic design. Its focus is on the evaluative thinking that should underpin the writing-up process: how to organise the key building blocks so that the audience can understand the assessment reached, why it matters, the evidence that supports it, and the implications for future decisions or action.

Who the guide is aimed at

In a nutshell, evaluators looking to introduce a more coherent and systematic approach to their analysis and reporting, whether they are early on in their evaluation career or have some experience but want a fresh perspective on how to go about it. It may also be useful to project managers, commissioners and others involved in making sense of evaluation reporting.

Be warned, though: if you are just starting out, this is not a primer. The guide assumes some familiarity with evaluation terminology, although I have included a glossary to show how I use certain terms throughout the guide.

The guide also assumes that you are working within the tradition of theory-based evaluation, with at least a basic understanding of what this entails.

It does not offer guidance on evaluation planning itself, or on how to build or refine a Theory of Change. A further reading list at the end of this guide is a good place to start if you want to know more about theory-based approaches, but it is not essential if you prefer to just dive into the five building blocks approach.

Nor does the guide favour one methodology over another, whether you use quantitative, qualitative or mixed-methods data collection.

The five building blocks apply equally well whichever approach you follow, and the guide does not attempt to help you choose between them.

Wider considerations

While this guide makes the case that the five building blocks provide a useful structure for analysis and reporting, they are not the whole of evaluation.

For example, sound practice can also depend on having clear criteria and standards for making assessments, paying attention to stakeholder perspectives, consideration of methodological strengths and limitations, appropriate handling of uncertainty, and credible reasoning about causation.

These issues underpin many of the judgements we make as evaluators, but they are not the primary focus here.

Rather than attempting to cover all aspects of evaluation practice, the guide concentrates on a narrower question: how to structure evaluative thinking and reasoning in a way that supports clear judgement and effective communication.

Guide structure

After introducing the building blocks in the next section, each block is described in further detail, including its purpose, where it fits within evaluative thinking and reasoning, and how its use differs (or not) between analysis and reporting tasks.

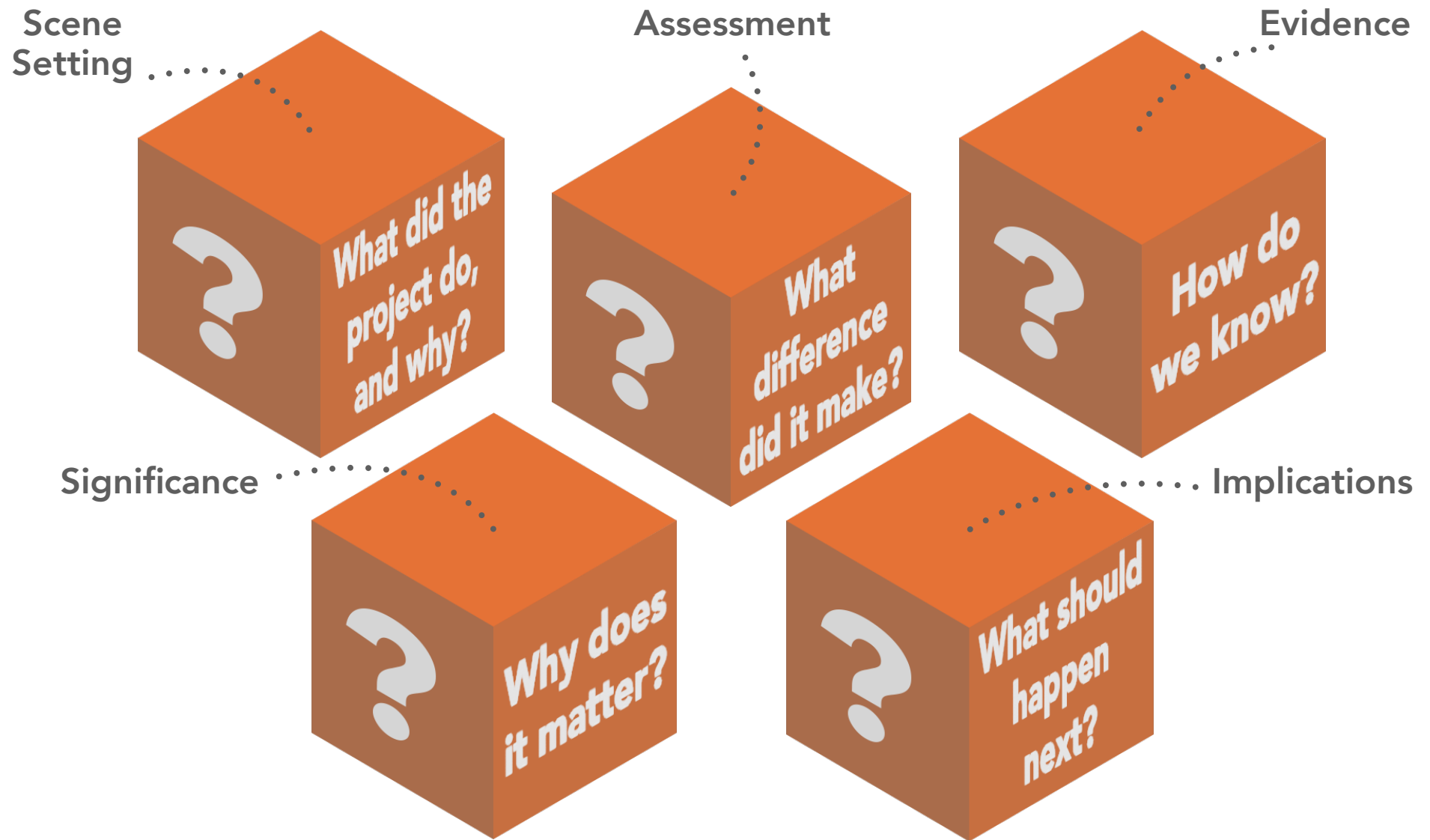
This is illustrated with examples drawn from two fictional projects: **Brighter Start Eastmere** and **Step Forward Now**.

To make these examples useful in their own right, some basic design and planning documentation for both imaginary projects is provided in the Annex (including simplified descriptions of their Theory of Change, underlying assumptions, key evaluation questions, data collection plans and outcome rubrics).

The guide also includes examples of published reports that you can download and read to see the building blocks in action.

This shows how the approach is flexible enough for use in a wide variety of reporting structures and styles, and for different purposes and audiences, and is equally applicable to evaluations conducted while a project is ongoing or at its conclusion.

Introducing the five building blocks



The five building blocks are intended to guide the way you think about, and write up, your evaluation evidence and judgements.

All five are necessary to provide a fully rounded account of your findings and how they should be acted upon.

Scene Setting

When considering Scene Setting, you are identifying relevant background information about the project you are evaluating, and why it was designed this way. This is project background, not evaluation methodology: what was delivered (including project outputs), the problem it addressed, and any information an analyst or report reader needs to orientate themselves.

Evidence

This is the data and information that supports an assessment. It involves gathering and interrogating data against pre-defined indicators and evaluation questions, including an active search for evidence that challenges any existing beliefs about project impact.

Assessment

This is when you arrive at an overall judgement about an outcome, project strand, or the project as a whole. Where a rubric exists to guide your assessment in a structured and consistent way, this is the point at which evidence is matched against it to reach a judgement based on agreed standards.

Significance

This is where you consider why the assessment matters to a particular audience. Not every finding is important, and this building block helps identify which are useful for informing a key decision about the project or which challenge an assumption, and for whom. As an essential building block, it should not be confused with the idea of statistical significance (which has a technical meaning in research and evaluation circles).

Implications

This is where you consider what your assessment means for future learning, decisions and action. It is where you look ahead to what should happen next, including updating your assumptions underpinning the project's Theory of Change.

Flexible thinking

While the blocks are presented separately for clarity, evaluative thinking and reasoning often develops through repeated movement between **Scene Setting**, **Evidence**, **Assessment**, **Significance** and **Implications**.

For example, new **Evidence** may alter how the **Scene Setting** is understood, and emerging **Assessments** may prompt a closer examination of the **Evidence**.

Similarly, evaluative thinking and reasoning often advances as your attention switches back and forth between analysis and reporting. But for simplicity, the guide assumes that report writing occurs

after analysis has concluded (whereas in reality they often happen at the same time, or overlap to a significant degree). This is not intended to imply that this is how things should always work.

The building blocks are therefore best thought of as elements of evaluative thinking and reasoning rather than stages in a linear process. They are presented individually to make the framework easier to understand and apply, while recognising that evaluation rarely proceeds in such an orderly fashion.

Reporting styles

Evaluation reports commonly focus on data collection and interpretation, beginning with an introduction and background to the evaluation (including the methods used), followed by a lengthy walk through the evidence, arriving at an assessment and recommendations later on. This **evidence-led approach** suits reports whose primary purpose is to document a body of evidence for the record.

As an alternative, this guide also considers the option of a **judgement-led approach**. This involves stating the assessment early, then justifying it, so that busy readers get the main points before the detail. This structure is broadly similar to the BLUF (Bottom Line Up Front) format used in some types of military and business writing.

Both reporting styles have their place. Evidence-led reports may be more appropriate for accountability or academic purposes.

In contrast, judgement-led reports may be more effective for executive decision making and are worth considering rather than defaulting to an evidence-led report structure out of habit.

Important note:

Whether you choose to follow an evidence-led or judgement-led format, the building blocks are not intended to function as section headings in an evaluation report.

They should help to shape your evaluation story, but any credible and useful report will require additional content (for example, details of the evaluation methods used) and a structure that suits its purpose and intended audience.

To put this more concretely, the following two pages show how the building blocks can be used to frame an evaluation report using either an evidence-led or judgement-led approach.

By illustrating this with traditional report section headings, I hope to make the underlying logic easier to grasp and ready to apply to your own reporting.

Evidence-led report example



This is how building blocks can be used to frame the narrative in an **evidence-led** report using traditional section headings

EXECUTIVE SUMMARY

Use all five building blocks to create an overview of the project, evaluation findings & recommendations. This ensures you provide readers with every essential piece of information they need in summary form

INTRODUCTION

Use **Scene Setting** to help frame your account of the project (including project outputs) and your approach to the evaluation (e.g. your key evaluation questions)

METHODS

Building blocks are not needed for this section, which provides an account of the data collection methods used, and any limitations that readers should be aware of when weighing up your findings

FINDINGS

Use **Scene Setting** to set out the relevant evaluation questions you are addressing in this section.

Use **Evidence** to provide the data and information upon which your assessment is based.

Use **Assessment** to describe your judgement (at the level of individual outcomes, and/or the impact of the project as a whole)

DISCUSSION & RECOMMENDATIONS

Use **Significance** to tie all the threads of evidence together and show how they are relevant to the underlying challenge or problem your project was addressing.

Use **Implications** to highlight what should happen next, and actions for key decision makers

Judgement-led report example



This is how building blocks can be used to frame the narrative in a judgement-led report using traditional section headings

EXECUTIVE SUMMARY

Use all five building blocks to create an overview of the project, evaluation findings & recommendations. This ensures you provide readers with every essential piece of information they need in summary form

INTRODUCTION

Use **Scene Setting** to help frame your account of the project (including project outputs) and your approach to the evaluation (e.g. your key evaluation questions). You should also include brief information about the methods used, to give readers confidence in what follows (further detail about your methods should be given in a technical annex)

FINDINGS

This section can be written at the level of the project as a whole, or the account can describe your findings outcome by outcome. Use **Scene Setting** to set out the relevant evaluation questions you are addressing in this section. Use **Assessment** to describe your judgement. Use **Significance** to help explain why this matters to particular readers/ stakeholders. Use **Evidence** to provide minimum necessary detail about the data and information underpinning your assessment. Use **Implications** to highlight what should happen next, and what action this entails for key decision makers

DISCUSSION & RECOMMENDATIONS

Use **Significance** to tie the evidence together and show how it is relevant to the underlying issue your project was addressing. Summarise the main points from your findings section (**Evidence**), concluding with an overall **Assessment** that balances all the evidence and makes clear recommendations for future decisions and action (**Implications**)

ANNEX: METHODS

Building blocks are not needed for this section, which provides an account of the data collection methods used, and any limitations that readers should be aware of when weighing up your findings

Deciding which report structure to use

When deciding which structure to adopt, whether evidence-led or judgement-led, this guide can help with conversations between evaluators and commissioners about what kind of report they actually want.

The five building blocks provide a practical, shared language for agreeing whether the report should be mainly evidence-led or judgement-led, based on its purpose and intended audience.

This conversation is best had early on, ideally when scoping the evaluation or agreeing reporting expectations, rather than after drafting has begun.

Tailoring your reports

There is plenty of scope to use the building blocks in different ways within a traditional report structure. Provided your report covers the main considerations associated with all five building blocks, it is not necessary to stick rigidly to any particular template.

For example, in the evaluation of projects with a large number of outcomes it may be preferable to provide details about the **Significance** and **Implications** of particular findings for each individual outcome within the Findings section.

That way, readers can take on board any relevant **Scene Setting**, your **Assessment**, the **Evidence** and its **Significance** and **Implications** at an individual outcome level, rather than receiving this information in one go for the project as a whole (you can

always draw all these elements together for consideration at the project level in a concluding Discussion and Recommendations section).

The modular nature of the five building blocks means you can combine them in whichever ways serve your reporting preference and the needs of your intended audience(s).

Bringing theory-based evaluation to life

One of the most powerful benefits of the five building blocks model is how it can strengthen theory-based evaluation.

Evaluators and project teams often develop Theories of Change or logic models at the outset but then struggle to make these frameworks meaningful in their write-ups. The underlying project logic appears only in an appendix or methodology section, disconnected from the analysis.

The five building blocks model prompts anyone delivering an evaluation to engage with project theory throughout the analysis and reporting stages, particularly when considering **Significance** and **Implications**.

Significance requires examining assumptions

When you ask "why does this assessment matter?" you should return to the underlying project logic. If an outcome was achieved or not achieved, what does this say about the relevant assumptions that underpinned the Theory of Change?

For example:

- If outcome X was achieved but outcome Y was not, what does this reveal about the assumed link between them?
- If an intervention worked in some settings but not others, what does this tell us about enabling conditions?

The part of your report dealing with **Significance** becomes a space to make these theoretical insights explicit and relevant to decision makers (whether they are part of the project delivery team, or a funder or commissioner).

Implications require acting on project logic

Following on from the above, when considering your findings' **Implications**, you should naturally ask: given what we have learned about how this project actually worked (or did not work), what does this mean for the underlying theory?

This might involve articulating what we now understand about "how change happens" and making recommendations about how to update and refine the theory, or what to consider when planning similar projects in future.

Rather than treating the Theory of Change as a static framework for organising findings, this approach supports a key principle of theory-based evaluation by encouraging evaluators to see it as a working model that the evaluation should test and refine.

The beauty of this approach is that it translates theory-related insights into language that practitioners can use.

Instead of an abstract discussion of "revised causal pathways," you are explaining why a particular finding matters (because it reveals something about how change actually happens) and what should be done differently next time (based on this improved understanding).

This means theory-based evaluation becomes genuinely useful for informing future decision making.

A note on outputs

Outputs are the immediate products of project delivery (e.g. sessions delivered, people reached or resources produced). In analysis and reporting, they are first considered as part of **Scene Setting**, because they clarify what was intended to happen and what actually happened in practice.

However, outputs should not be treated simply as an administrative count of activity. They can be judged in their own right, by asking whether delivery was sufficient in scale, reached the intended people, was delivered consistently and to an appropriate standard, and was accessible and equitable.

In this case, outputs can also be considered as **Evidence**, used to support a distinct **Assessment** of implementation or delivery quality (as part of a process evaluation), alongside judgements about whether intended outcomes were achieved (impact evaluation).



Scene Setting

How Scene Setting contributes to ANALYSIS	How Scene Setting contributes to REPORTING
Purpose Establishes the frame of reference for interpreting the evidence, so analysis has something to be judged against rather than drifting towards whatever is most visible in the data. Features Draws on the project's Theory of Change, data collection plan, indicators and output measures, rubric, and key evaluation questions defined in advance. Questions you should consider: • Am I approaching the data with a ready-made frame of reference, or is this an open-ended enquiry? • What was this project trying to achieve, and by what assumed mechanism(s)? • What indicators and evaluation questions were set out in advance? • What outputs was the project expected to produce, at what scale, for whom, and what do we know about what was actually delivered? • If there is a rubric, what does success look like? If there is no rubric, how should I define what success looks like?	Purpose Primarily background, context and project description, orientating the reader to what is being evaluated and why it exists, before they encounter any judgement or outcomes evidence. Features Project background, not methodology: the problem addressed, intended beneficiaries, delivery model and the principal outputs actually delivered. Questions you should consider: • What does the reader need to know to make sense of the project and my assessment? • Have I kept this to background and descriptive information rather than moving into methodology or findings? • Would a reader unfamiliar with the project understand what is being discussed from this content alone?

Scene Setting is the building block that answers some deceptively simple questions: why does the project exist, why has it been designed this way, and what, exactly, is being evaluated?

It sounds obvious, but this building block is commonly rushed or under-developed, which risks creating sloppy analysis and confusing reports.

Think of **Scene Setting** as essential context. It tells you (as the evaluator), and later your reader, what a project was trying to do, what was delivered and for whom, and how it expected to achieve its goals, before getting into whether it succeeded or not.

Before beginning your data analysis, good practice involves returning to a pre-defined evaluation framework rather than approaching the data cold. This framework typically includes:

- The Theory of Change or logic model underpinning the project
- The data collection plan that determined what evidence was gathered, and how
- The indicators set out in advance against which outcomes are being judged, along with any output measures (and targets) for quantifying what was delivered
- Any evaluation rubric used to define what different levels (standards) of performance look like across relevant criteria
- The key evaluation questions the evaluation was designed to answer

Together, these establish the frame of reference for the analysis: what the project was trying to achieve, via what assumed

mechanism, and the definition of what success looks like before any data was collected.

Without this framing, analysis has no reference point against which to interpret what is found, and risks being driven by whatever happens to be most visible in the data rather than by what the evaluation set out to determine.

If Scene Setting is incomplete

In real world evaluation, this ideal framework is not always in place.

For example, some projects may not have a full Theory of Change or similar, or the commissioning brief may arrive with broad or vague evaluation questions.

Where the ideal framework is missing, you can establish a minimum viable **Scene Setting** before starting on the data, by answering the following:

- What was the project trying to achieve?
- What was it expected to deliver (and were any targets set at the start)?
- Who was it intended to benefit?
- What outcomes matter most?
- What evidence is available on those outcomes?
- What would count, broadly, as success, partial success, or limited success?

This lighter framework is less robust than a fully designed evaluation framework, but it is still better than approaching the data with no reference point at all.

Why Scene Setting comes first, in both analysis and reporting

Scene Setting is unusual among the five building blocks because it holds the same position whether you are analysing data or writing up your findings. This is because everything else depends on it.

What needs to be considered when framing Scene Setting

The easiest way to grasp **Scene Setting** is to understand what it deliberately leaves out. **Scene Setting** is not methodology, evidence or judgement. It is essential background and project description only.

A useful test is to think about if you removed the parts of your reporting that deal with **Scene Setting**, would the reader still understand why the project existed, what it was trying to achieve and what was delivered? If not, some crucial details are missing.

Example: Brighter Start Eastmere literacy project

Imagine a fictional evaluation of Brighter Start Eastmere, a two-year project run by a charity to improve reading ability among seven- to nine-year-olds in primary schools in a town called Eastmere.

Weak way of framing Scene Setting in reporting:

"Brighter Start Eastmere was delivered over two years and aimed to help improve children's reading ages."

This is factually correct, but provides none of the specific detail readers need about the problem it was designed to address, the delivery model, target age group and underlying mechanism to bring about the intended change.

Another example:

"Brighter Start Eastmere was evaluated using a mixed-methods approach involving pre- and post-intervention reading assessments across six schools."

This is methodology, not **Scene Setting**. A reader still has no idea what Brighter Start Eastmere actually does or why it exists.

Strong way of framing Scene Setting in reporting:

"Brighter Start Eastmere placed trained reading volunteers in six primary schools in Eastmere for two hours a week over two academic years, working with children identified by teachers as reading below their expected age level. The project was designed to address a persistent gap: national data showed nearly a third of children in the area were behind on reading benchmarks by age seven, with the gap widening rather than narrowing through primary school."

Brighter Start Eastmere's underlying assumption was that regular one-to-one reading support, delivered consistently over time,

could close this gap before it hardened into a longer-term disadvantage.”

Notice that the strong framing of **Scene Setting** says what was delivered (volunteer reading support, two hours a week, two schools years), to whom (children reading below expected level), why it existed (a documented, worsening gap), and the assumed mechanism (consistent one-to-one support closes the gap).

None of this tells the reader whether Brighter Start Eastmere worked (that comes later) but without it, nothing that follows would make sense.

How the framing of Scene Setting changes between analysis and reporting

Even with the same underlying facts, what an analyst does with **Scene Setting** differs from what a report shows the reader.

In analysis, working through **Scene Setting** for Brighter Start Eastmere means actively retrieving and checking documents that may not have been touched since the project was designed: was there a formal Theory of Change setting out the assumed mechanism? Were reading-age benchmarks specifically defined in the data collection plan as the indicator of success? Was there a rubric distinguishing "some progress," "closed the gap," and "no measurable change"?

Doing this before looking at a single results spreadsheet stops the analyst from later chasing whichever numbers look most impressive.

In reporting, **Scene Setting** becomes a short, readable orientation for someone who may never see the Theory of Change document at all (e.g. a school governor, a funder's trustee, a journalist etc.). The reporting version of **Scene Setting** for Brighter Start Eastmere would not mention the data collection plan or indicators explicitly. It would simply set the scene, as in the example above, so the reader is ready for the judgement that follows later.

Example: Step Forward Now scheme

A second example helps show how **Scene Setting** can be applied to different subject matter. Step Forward Now is a fictional six-month project helping long-term unemployed adults into work through coaching and short vocational courses.

Good reporting **Scene Setting** would state that:

“Step Forward Now worked with adults unemployed for over 12 months, referred by local job centres. It combined weekly coaching with access to short, funded vocational courses in care work and logistics chosen to address local skills gaps. It was designed on the assumption that the primary barrier to employment for this group is a skills deficit relative to available roles.”

This **Scene Setting** helps to distinguish the project, in the reader's mind, from a purely training-focused scheme.

A reader now understands not just what Step Forward Now did, but what problem it was specifically designed to solve, which is exactly what they need before encountering any judgement about whether that was achieved.

In short

Scene Setting is the foundation that explains what a project was intended to achieve, for whom, and how. In analysis, it provides the frame of reference for interpreting evidence, while in reporting, it provides the background that allows readers to make sense of the evaluation's findings.

The Brighter Start Eastmere and Step Forward Now examples show that strong **Scene Setting** in reporting typically includes details about:

- the intervention as it was delivered
- the target beneficiaries
- the problem being addressed
- the project rationale
- the assumed mechanism of change

Importantly, it does not say whether the project succeeded. That judgement comes later.

Top tips

1. Start with key project details

Before analysing results, remind yourself what the project was trying to achieve, for whom, how it was expected to create change, and how you would know if it succeeded.

2. Separate Scene Setting from methodology

Scene Setting explains the project and its purpose, while methodology explains how it was evaluated. Do not conflate the two.

3. In reports, use Scene Setting to provide essential context

Readers need to understand what was delivered, who benefited, what problem was being addressed, and why the project existed before they can make sense of findings.

4. Explain the underlying logic

Include the project's key assumption about how change was expected to happen, because this helps others to understand the basis for later judgements.

5. Keep Scene Setting factual, neutral and concise

Scene Setting in reporting should describe the project, not whether it succeeded. Save evidence and judgement for later sections. Keep it concise by providing essential details only.

How EVIDENCE contributes to ANALYSIS	How EVIDENCE contributes to REPORTING
Purpose Forms the analytical base before a judgement is made, as a deliberate safeguard against anchoring bias (the tendency to rely too much on the first piece of information found). Features Gathering and interrogating data against the indicators and questions set out in Scene Setting, with an active search for evidence that challenges an early hunch and not just confirms it. Questions you should consider: • What does the data show, across all relevant sources and indicators? • How robust and credible is the evidence across different sources? • Have I actively looked for evidence that would challenge a hunch? • If evidence sources conflict, have I properly triangulated and weighed the findings rather than just picking the most convenient one? • Am I ready to form a judgement, or is there more to interrogate first?	Purpose Demonstrates the basis on which the stated judgement was reached, giving the reader confidence in it. Features Evidence presented with a clear purpose which is comprehensive but proportionate and tied explicitly to your assessment. Questions you should consider: • Does each piece of evidence clearly contribute to the assessment? • Have I kept my description of the evidence proportionate rather than exhaustive? • Would a reader who already knows my conclusion think this section builds their confidence in my findings rather than overwhelming them with detail?

Evidence is the building block where claims about what a project will achieve stop being assumptions or intentions and start being tested against what really happened.

The job of **Evidence** is not to prove a point you have already decided on, it is to find out, honestly, which data support an evaluation judgement, including the parts of your evidence base that are inconvenient.

Why Evidence's position changes with the task

Evidence's position in respect of other building blocks moves depending on whether you are analysing or reporting.

Understanding why reveals something important about what

Evidence is actually for.

- In analysis, consideration of **Evidence** comes before **Assessment**, straight after **Scene Setting**. This ordering exists because of a tendency called anchoring: once you have formed a view, you unconsciously start noticing evidence that confirms it while skimming past evidence that does not (leading to confirmation bias). Working through **Evidence** before concluding with a judgement forces you to look at everything the data shows before deciding what it means, rather than cherry-picking data to support a conclusion you already reached.
- In reporting, the reader is not at risk of anchoring, because they are not the one forming the judgement from scratch. Here **Evidence's** job shifts from protecting against premature

conclusions to demonstrating that the stated conclusion is warranted.

What stays constant across both analysis and reporting is the underlying discipline of handling and interpreting data. **Evidence** must be gathered and presented honestly, including anything that complicates the picture.

What needs to be considered when framing your Evidence

Evidence means the data itself, whether that is routine monitoring information, survey responses, interview transcripts, attendance records, or whatever other sources the evaluation draws on.

A good approach to **Evidence** involves more than reporting what the data shows. It requires actively hunting for evidence that might contradict an emerging or stated view, and being honest when different sources point in different directions rather than quietly favouring the one that tells a tidier story.

As a starting point, you can consider the following questions about your evidence base:

- What does each source say about the outcome or evaluation question?
- Which sources of evidence point in the same direction?
- Which evidence conflicts or raises doubt?
- Are there major gaps in the evidence base?
- Am I giving too much weight to the most vivid or recent data?

In both analysis and reporting, it is useful to ask whether a colleague or reader could see how a different judgement might have been reached if they consulted only the evidence you have gathered.

If your **Evidence** only seems to support one conclusion, it is worth asking whether contradictory data was actually looked for, or simply not found because it does not exist.

Example: Brighter Start Eastmere's reading data

Suppose the fictional Brighter Start Eastmere evaluation gathered three main evidence sources: reading age assessments before and after the project, teacher observations and attendance records.

Weak way of framing Evidence in reporting:

"Most children who took part in Brighter Start Eastmere made excellent progress, showing the project was a clear success."

This mixes and conflates **Evidence** and **Assessment**. It does not say what the data actually showed, does not quantify "most" or "excellent," and does not mention any data source that might complicate the picture.

Strong way of framing Evidence in reporting:

"Reading age assessments showed that 68% of participating children gained at least six months of reading age over the two-year project, against an expected gain of four months for their peer group."

However, this finding masked considerable variation: children with attendance above 80% gained an average of nine months, while those with attendance below 50% (roughly a fifth of participants) showed gains barely different from non-participants.

Teacher observations broadly corroborated the quantitative gains for regular attenders, but several teachers also noted that reading confidence, not just reading age, seemed to improve even among some lower attending children, which is a dimension the reading age assessments did not capture."

Notice how this version gives actual figures from the data record, it surfaces a pattern that complicates a simple "it worked" story (the attendance effect), and it flags a gap between what different sources measured (confidence versus reading age) rather than smoothing that gap over.

Example: Step Forward Now's employment data

For Step Forward Now, the fictional employment support scheme, imagine three data sources: job entry records six months after the project, coach case notes and a short participant survey.

A weak way of framing Evidence here might say: *"Participants overwhelmingly found Step Forward Now helpful and most moved into work."*

A stronger way would instead report that:

“Two-fifths (42%) of participants were in work six months after finishing the project, compared with a national benchmark of around 20% for adults with a similarly long unemployment history. However, job entry was markedly lower among participants who had been unemployed for more than three years going in.

Coach case notes show that confidence building was the most frequently cited benefit even among those who had not yet found work, and the participant survey showed high satisfaction with coaching but notably lower satisfaction with the vocational courses on offer.”

On this basis, a reader now has a genuinely nuanced, evidenced picture rather than a single flattering headline, including an important signal that the coaching element may be contributing more than the vocational courses.

How the framing of Evidence differs in practice between analysis and reporting

In analysis, working through **Evidence** means going wide before narrowing. This entails pulling every relevant data source, checking sources against each other (triangulating), and specifically asking "what would tell me this didn't work?" rather than only "what tells me it did?"

This is deliberately laborious, because its purpose is to protect the judgement that comes next from being formed too quickly.

For the Brighter Start Eastmere example, the reporting version of **Evidence** might lead with the headline reading age figures and the attendance pattern, since those are most relevant to an assessment, leaving more granular detail to an evidence appendix (e.g. school-by-school breakdowns).

Crucially, though, the report should still surface the attendance effect and the confidence/reading age gap, not quietly drop them because they complicate any judgement. A credible **Evidence** section in a report should always acknowledge complexity.

What to do when your Evidence has gaps or obvious weaknesses

Data gathering rarely runs to plan, and any number of factors, which may be beyond your control, can affect the quality and comprehensiveness of your evidence base (e.g. low survey response rates, short fieldwork windows with limited opportunities for qualitative data collection, incomplete project records and monitoring data etc.).

In these circumstances, specific limitations or gaps should be flagged in your reporting, showing how you have factored them into your handling of the **Evidence**, your **Assessment** and your consideration of **Significance** and **Implications**.

In short

Evidence is not there to prove what you already believe; it is there to test and challenge potential conclusions before you

finalise them, and then justify those conclusions when you report them.

In both cases, credible evaluation requires presenting the full picture, including findings that complicate your preferred narrative along with any gaps or weaknesses in the evidence base.

Top tips

1. Look at evidence before forming a firm judgement

Collate and interrogate the full range of data before deciding what it means. This helps avoid anchoring and confirmation bias and prevents conclusions being formed too early.

2. Actively seek contradictory evidence

Do not just look for data that supports your emerging view. Consider what evidence would suggest the project did not work and actively search for it.

3. Triangulate multiple sources

Compare findings across different datasets, surveys, interviews, observations and records. Where sources disagree, investigate why rather than selecting the most convenient finding.

4. Give appropriate weight to evidence

Avoid being unduly influenced by the most recent, vivid or memorable information, and consider the overall strength and credibility of the evidence base.

5. Use evidence to support your assessment

Every piece of evidence reported should clearly contribute to explaining and justifying the conclusion you reached.

6. Be comprehensive but proportionate

Provide enough evidence to build confidence in the findings without overwhelming the reader with unnecessary detail (detailed data can be placed in appendices if needed).

7. Always highlight evidence gaps or limitations

If you lack sufficient high quality evidence to make an **Assessment** with any degree confidence, say so in the report so that readers can understand the basis of your judgement and weigh it for themselves.

How **ASSESSMENT** contributes to ANALYSIS

Purpose

Commits to a judgement once the evidence has been properly explored, making the conclusion evidence-based.

Features

Considers all evidence in the round, weighing it to form a balanced determination in answer to a specific evaluation question. Where a rubric exists, evidence is matched against it to reach a judgement. May be provisional and revised through the Evidence/Assessment loop as other sources are triangulated.

Questions you should consider:

- What is the clearest, fair judgement supported by the evidence?
- Has the evidence been explored thoroughly enough to support a judgement, or am I forming this too early?
- Where a rubric exists, how does the evidence rate against it?
- Is my judgement specific enough to be tested further, or is it too vague to commit to?
- Does new evidence require me to revise a provisional judgement?
- Could another evaluator reach a different judgement using the same evidence base? If so, how and why?

How **ASSESSMENT** contributes to REPORTING

Purpose

States the judgement clearly, giving the reader a firm basis for interpreting everything that follows.

Features

Uses clear, unhedged language unless qualifying the judgement is justified.

Questions you should consider:

- Have I stated my judgement clearly and directly, without unnecessary hedging?
- Am I resisting stating this upfront out of discomfort (because it might disappoint a client, or invite pushback, for example) rather than out of genuine uncertainty?
- Would a time poor reader understand my conclusion from this section of the report alone?

Assessment is the building block where an evaluation starts drawing conclusions about what it found.

During analysis, it is the point at which everything considered when thinking about **Scene Setting** and **Evidence** is converted into a clear judgement (e.g. did change occur as planned/ did this project work, to what extent, and against what standard?).

Assessment requires genuine commitment. A vague or heavily qualified assessment risks leaving the analyst or report reader unable to test or act on it.

Why Assessment demands clarity

In analysis, **Assessment** comes after **Evidence** has been properly explored, to avoid the risk of anchoring. It may be provisional, revisited as further evidence is triangulated, before it is settled enough to carry forward into your reporting.

In judgement-led reporting, **Assessment** typically comes early (i.e. straight after **Scene Setting**, before **Evidence**) presenting the verdict first, as a basis for interpreting what follows.

In practice, the test of a good **Assessment** is the same whether you are making it privately or stating it publicly. Would a reasonable person looking at the same evidence be able to agree or disagree with what you have said, or is it too vague to pin down?

What a strong Assessment looks like

A weak way of framing an **Assessment** in reporting hedges without good reason, for example: *"Brighter Start Eastmere showed some positive indicators in certain areas, although results varied."*

This sounds careful, but it commits to nothing a reader could disagree with, test or act on, which limits its usefulness.

A strong way of framing an **Assessment** in reporting presents a specific verdict, calibrated to the evidence:

"Brighter Start Eastmere met its core aim of accelerating reading progress for the majority of participants who attended consistently, but fell short of delivering that benefit to all because a sizeable minority did not attend consistently enough to benefit."

Forming an **Assessment**, whether during analysis or stating it in a report, means resisting the temptation to round up to "Brighter Start Eastmere worked" (too generous, given variation due to attendance consistency) or round down to "Brighter Start Eastmere failed to deliver consistent results" (too harsh, given that most children who attended regularly saw gains well above the benchmark).

Where Brighter Start Eastmere had "Reduced gap between participating children's reading age and chronological age" as an outcome,

an honest rating using the standards in the associated rubric would be something like "partially met, with attendance identified as the primary limiting factor."

Example: Step Forward Now

In the case of Step Forward Now, **Evidence** showed that job entry comfortably exceeded a 20% benchmark overall, but this finding alone masked the contribution made by coaching (compared with the effect of vocational courses) and weaker outcomes for those unemployed longest.

This example illustrates a different facet of **Assessment**, namely recognising when more than one distinct judgement is warranted, rather than collapsing everything into a single verdict.

If the key evaluation question was simply "did Step Forward Now improve job entry relative to the benchmark level?", the evidence supports a clear, positive **Assessment**.

But if the evaluation was also tasked with judging whether the project's design operated as intended, this would require a further judgement, and the evidence there is more qualified. The outcome was achieved, but not definitively through the mechanism the design assumed (i.e. vocational courses were expected to deliver most benefit, supported by coaching).

The discipline **Assessment** demands here, in both analysis and reporting, is keeping these two judgements distinct rather than merging them into a single, blurred statement that either

undersells the strong headline result or overstates confidence in the project's design logic.

In analysis, this means considering both key evaluation questions before settling on two separate judgements. In reporting, it means giving the reader both judgements in sequence, clearly distinguished:

"Step Forward Now substantially outperformed the national benchmark for job entry, achieving its core job entry outcome. However, the evidence does not support the project's original assumption that vocational courses were the primary driver of this success."

In short

Assessment involves moving from examining evidence to making clear, evidence-based judgements, and the same discipline is required whether you are analysing findings privately or reporting them publicly.

A good **Assessment** is a clear and specific judgement that accurately reflects the full range of evidence, answers the key evaluation questions, and is stated confidently enough for others to test and act upon.

Top tips

1. Use Scene Setting as your frame of reference

Judge findings against the project's intended outcomes, indicators, evaluation questions, and success criteria, not just against whatever patterns stand out in the data.

2. Do not form a judgement too early

Explore the evidence thoroughly before reaching a verdict, as early conclusions can create bias by encouraging you to interpret later evidence in ways that support your first impression.

3. Base conclusions on all the evidence

Consider evidence "in the round" rather than relying on a single finding, and weigh different sources and triangulate evidence before settling on a conclusion.

4. During analysis, be willing to revise your judgement

Treat early assessments as provisional if more evidence is still being gathered or collated. You should update conclusions when new evidence genuinely changes the picture.

5. Use rubrics where available

Judge performance against agreed standards rather than personal opinion or unstated criteria.

6. Always be open and transparent, whatever the verdict

Be sure to weigh evidence honestly, even when the conclusion may be unfavourable. This can be just as useful for learning and decision-making.

7. Avoid vague or heavily hedged language

Strong assessments state the verdict directly unless genuine uncertainty requires some qualification.

8. Be balanced, not overly positive or negative

Resist the temptation to oversimplify results into "it worked" or "it failed". Aim to reflect both achievements and limitations supported by the evidence.

9. Keep different judgements separate

Different evaluation questions may require different conclusions, and you should avoid merging distinct findings into a single headline statement.

How SIGNIFICANCE contributes to ANALYSIS	How SIGNIFICANCE contributes to REPORTING
Purpose Identifies which findings would genuinely inform a decision about the project or future activity, or which challenge an assumption in the underlying Theory of Change, so time is not spent writing up findings of little or no consequence. Features Requires asking “significant to whom and to what end?”, since funders, delivery partners and beneficiaries may weigh findings differently. Questions you should consider: • Which findings are most consequential in addressing key evaluation questions and informing future decisions and action? • To whom is this significant, and would different stakeholders weigh this finding differently? • Would ignoring a particular finding create a missed opportunity? • Am I about to spend time writing up something merely interesting rather than something that matters?	Purpose Explicitly answers “so what?” for the intended audience reading the report, translating an Assessment into terms that matter to them. Features Connects the Assessment to what is at stake for a particular readership, not a generic statement of importance. Questions you should consider: • Why should this particular audience care about this judgement? • What is at stake for them if the lesson or insight goes unheeded? • Have I translated this into terms that matter to a reader, rather than assuming the significance is self-evident?

Consideration of **Significance** stops an evaluation from treating every finding as equally important. It answers a question that is easy to skip past: of everything the **Evidence** turned up, what actually matters enough to shape a judgement or to be reported to share learning, and to whom?

To be clear: **Significance** is not a restatement of the assessment in stronger language. And in this context it has nothing to do with the technical concept of statistical significance. Rather, it is a distinct, deliberate step in the evaluative reasoning process that asks "So what?", and insists on answering it for a specific audience, rather than assuming the answer is obvious and equally applicable to all stakeholders.

Why Significance requires identifying an audience

Significance only does useful work once you have specified whose significance you mean. A finding that stands out to a funder might be a non-issue for a frontline delivery team, and a pattern of findings about participant satisfaction might barely register with a commissioner focused only on cost-effectiveness.

This is true whether **Significance** is being worked through during analysis or when reporting. In both cases, this way of thinking risks falling short if the audience is left unspecified.

What Significance is not

Significance is easily confused with **Scene Setting** (which explains the project's background), and **Assessment** (which states the judgement itself). It is related to both, but distinct from them.

When you take time to consider **Significance** during your write-up, you are translating "this is what we found" into "this is why you should care."

Although they are treated separately for the sake of clearer explanation, thinking about **Significance** goes hand-in-hand with the consideration you should give to **Implications** (as described in the next section).

Example: Brighter Start Eastmere

Brighter Start Eastmere's evaluation **Evidence** showed strong reading age gains overall but a marked drop-off for children with attendance below 50% of the sessions.

In analysis, working through **Significance** means asking who would benefit from knowing about this attendance pattern, before deciding whether it is worth building into the report.

The project's own delivery team would likely find the attendance drop-off valuable to know, since it points directly at something they can act on (e.g. whether transport, scheduling or family circumstances are preventing consistent attendance).

An evaluator who stops in their evaluative reasoning at "children mostly did well", and does not flag the attendance pattern as significant to delivery staff, has missed the finding most likely to improve delivery for better outcomes.

In reporting, **Significance** means writing this out explicitly rather than assuming the reader will draw the same connection. For a delivery team readership, the report might say something like:

"The attendance threshold identified here should be treated as an operational priority. It suggests the current level of support is not enough on its own for around a fifth of participants, regardless of project quality. This matters because outcomes could be improved for a larger proportion of participants with better support for regular, consistent attendance."

Example: Step Forward Now

For Step Forward Now, recall that **Evidence** showed job entry was markedly lower for participants unemployed more than three years, and that coaching consistently outperformed the vocational courses in terms of participant satisfaction, even among those who had not found work.

In analysis, **Significance** here means resisting the temptation to treat both findings as equally reportable. The finding about long-term unemployment duration might be significant mainly as a caveat on the headline job entry figure. This is interesting, but not necessarily something the project can change through delivery.

The finding about coaching outperforming courses is significant in a more obvious way, as it directly challenges an assumption built into the project's design (that vocational courses were a core driver of employment outcomes) and points to a concrete decision (i.e. whether to shift investment from courses to coaching).

An evaluator considering **Significance** properly would prioritise carrying the second finding forward into their reporting, precisely because it is actionable, not just interesting.

In a report, this becomes a clearly stated case to the funder or commissioner deciding whether to renew or redesign Step Forward Now:

"The finding that coaching, not the vocational courses, appears to be driving participant confidence and engagement has direct implications for how any renewed funding should be allocated."

"Continuing to fund courses at current levels, without evidence they are the active ingredient in outcomes, risks misallocating resource that coaching capacity could use more effectively."

In short

Significance is the discipline of deciding what truly matters and why anyone should care about it. It prevents evaluators from treating all findings as equally important or relevant.

Consideration of **Significance** involves identifying which findings matter most, to whom they matter and why the findings should influence decisions or action, both during analysis and when communicating results.

Top tips

1. Ask “What actually matters and why?”

Focus on findings that could influence decisions, future activity or understanding of the Theory of Change.

2. Always ask “Significant to whom?”

Different stakeholders will value findings differently, so identify your key audience(s) before deciding what is significant. When reporting, explain why the finding matters to a particular readership, whether they are funders, commissioners, delivery staff, participants etc.

3. Avoid reporting things just because they are interesting

A useful test is whether ignoring a finding would create a missed opportunity or affect the basis and efficacy of future decisions.

4. When reporting, answer the reader's “So what?”

Make the significance of findings explicit rather than expecting readers to work it out for themselves. Explain what is at stake, and clarify the consequences of acting on, or ignoring, findings.

How IMPLICATIONS contributes to ANALYSIS	How IMPLICATIONS contributes to REPORTING
Purpose Examines what the findings reveal about the assumptions underpinning the project's Theory of Change, what this suggests about how change happened and whether the model needs to be revised. Connects your Assessment to future learning, decisions or action: what stakeholders should do differently. Features Treats the Theory of Change as a living hypothesis tested by the evaluation, not a static framework. Ensures findings have practical utility. Questions you should consider: • Which assumptions in the Theory of Change proved valid or not? • Did any unexpected causal pathways emerge? • What do I now understand about how change happens in this instance, that I did not before the analysis? • What lessons can genuinely be applied to this or future initiatives?	Purpose Translates your thinking about what should happen as a result of the findings into concrete, practical recommendations. Features Specific and actionable rather than vague and overly cautious. Questions you should consider: • What should happen next, specifically? • What is the most effective way to clearly, and directly, communicate WHO should do WHAT in future? • Are the recommendations concrete enough to act on, or do they need spelling out in more explicit detail (e.g. changing the framing from "consider these findings" to a specific action for a specific stakeholder tied to a specific finding)?

Implications is the building block that helps you turn a judgement into forward action.

Where **Significance** asks "Why does this matter, and to whom?", **Implications** asks a related but different question: "Given that this finding matters, what should you do about it?"

In other words, **Significance** explains what is at stake for your intended audience(s), while **Implications** specifies the action(s) that key decision makers should take.

Neglecting one or other of these crucial blocks will produce a report that either states the importance of a finding but never says what to do about it, or jumps straight to recommendations without ever establishing why anyone should care to act on them.

It is easy to see **Implications** as simply "the recommendations bit," but its real value goes wider than this. Consideration of **Implications** should involve asking what a finding reveals about the assumptions underlying how the project was expected to work, and what follows from that for future decisions, design or learning.

It can help to consider three main types of **Implication**:

- **Design Implications:** What should change in project design next time, including resource allocation?
- **Delivery Implications:** What should be implemented differently in practice?

- **Evidence Implications:** What should be monitored, evidenced or explored more carefully in future?

What weak framing of Implications looks like

Vagueness is the most common failure when considering and articulating **Implications**. In other words, a statement that sounds like an implication but commits to nothing. "Future projects should consider these findings..." is the classic example that acknowledges that something should happen without saying what, for whom or why.

In contrast, a strong way of framing **Implications** actually names the specific assumption tested, states what was learned about it, and ties that learning to a specific, actionable change.

Example: Brighter Start Eastmere

When considering Brighter Start Eastmere's evaluation **Evidence** and **Significance** together, it was established that attendance below 50% sharply reduced reading age gains, and that this mattered most directly to the delivery team responsible for project design.

In analysis, working through **Implications** means going back to Brighter Start Eastmere's original Theory of Change and asking what the attendance finding reveals about key assumptions.

The finding suggests that the assumption that consistent one-to-one support would close the reading gap was basically correct

(support does close the gap). However, the assumption was incomplete because it did not account for the fact that support only works if it is reliably delivered and received, and that something is currently preventing roughly a fifth of children from receiving it consistently.

In this example, the **Implication** is not "attendance matters" (that is **Significance**, already established). Instead, the key **Implication** is the more specific realisation that Brighter Start Eastmere's Theory of Change needs an additional, previously unstated assumption made explicit: that attendance barriers themselves need active management.

In reporting, this becomes a concrete, actionable statement rather than a vague gesture at "considering attendance."

A weak way of communicating this would be:

"Future delivery of Brighter Start Eastmere should keep attendance in mind."

A strong way of communicating it would be:

"Brighter Start Eastmere's next delivery cycle should introduce a light-touch attendance monitoring process, triggering a check-in call to families after two consecutive missed sessions, and should track whether this measurably narrows the gap between high- and low-attending participants in the next evaluation."

This tells the reader exactly what should change, who is responsible, and even how success would be checked next time, which are all hallmarks of a properly specific **Implication**.

Example: Step Forward Now

For Step Forward Now, **Significance** established that the coaching finding mattered specifically to decisions about how to allocate future investment.

In analysis, **Implications** here means asking what the finding reveals about Step Forward Now's founding assumption (that a lack of vocational skills was the primary barrier to work for long-term unemployed adults, and that funded courses were therefore the project's central lever).

The evidence suggests this assumption was at least partly wrong: confidence and structured support, delivered through coaching, appear to be doing more to drive positive outcomes than the courses themselves.

The **Implication** is not simply "coaching matters more than expected" (that is what a consideration of **Significance** has concluded) but the realisation that Step Forward Now's underlying model of how unemployment can be overcome for this group may need revising (moving from a 'skills deficit model' to a 'confidence building model'), which has consequences beyond this current project cohort.

In reporting the finding, this translates into a specific recommendation rather than a general nod to its importance.

A weak way of communicating it would be: *"Step Forward Now should think about the balance between coaching and courses."*

A stronger formulation would be something like:

"We recommend Step Forward Now pilot a reallocation of 20% of course budget into expanded coaching capacity for the next cohort, with job entry and confidence outcomes tracked separately by coaching intensity, to test directly whether coaching is the primary driver before committing to a permanent rebalancing of the model."

This gives a decision maker something they can act on and revisit in discussion with the delivery team, not just a general sense that coaching seems important.

In short

Implications bridges the gap between evaluation findings and action. It helps evaluators move beyond explaining why something matters to specifying what should be done next.

When considering **Implications**, you should translate evaluation findings into evidence-based changes to project design, delivery or learning, expressed as clear actions for identifiable decision makers rather than vague suggestions.

Top tips

1. For every finding you plan to take forward in your report, ask:

- What assumption did this finding test?
- What have we learned about how change happens?
- What should be done differently in terms of project design, delivery or evidence collection?

2. Look beyond the immediate finding

Consider what the evidence reveals about how change actually happened and whether unexpected causal pathways emerged.

3. Distinguish Significance from Implications

Do not stop at explaining why a finding matters; explore what that means for future decisions and actions.

4. Be specific and actionable

Recommendations should clearly state who should do what, rather than making general statements.

5. Define what success would look like

Where possible, include how the proposed action(s) should be monitored or tested.

Real world examples of the building blocks in action

You can see the building blocks in action in these published reports I've authored using the approach outlined in this guide (click report titles to download).

The blocks themselves are not named in any of the reports, but they underpin the evaluative thinking and reasoning implicit in the write-ups.

For each example, I've identified how the building blocks were used given the specific purpose of the evaluation, and where they can be found in the report so you can see for yourself.

Evidence-led report examples

[Young Audiences Content Fund Pilot: End-of-Term Evaluation \(December 2022, BFI\)](#)

This is an end-of-term evaluation of the YACF pilot. Its principal purpose is accountability, to provide DCMS, the BFI and other stakeholders with a comprehensive, independent assessment of whether the Fund delivered its intended social and economic value, was administered well and represented value for money.

That purpose shapes the balance of the five building blocks:

- **Scene Setting** is especially prominent as it explains the market failure the Fund was designed to address, the policy rationale for intervention, the Fund's delivery model and the

outcomes it was expected to achieve. *See Executive summary, About YACF, About the report, and the opening sections of Section 2: Impact evaluation*

- **Evidence** is extensive, covering supported projects, audience performance, content quality, additionality, diversity, plurality, innovation, employment, skills, regional activity, Fund processes and value for money. *See Impact headlines infographic, Executive summary, Sections 1 to 4 and supporting appendices*
- **Assessment** is equally prominent: the report judges each outcome in the pre-defined model and reaches an overall conclusion about whether the Fund generated the intended social and economic value. *See Executive summary and Sections 1 to 4*
- **Significance** explains why the findings matter for public policy. The central question is not simply whether projects were funded, but whether the Fund expanded the supply of content that the market would otherwise struggle to provide. *See Sections 2 to 5*
- **Implications** are strategic rather than operational. Because the pilot had ended, the report uses its findings to inform the policy decision about whether targeted contestable funding can provide a viable response to persistent gaps in public service content for young audiences in the UK. *See Section 5: Concluding discussion*

Film for Learning: Year Three (October 2022, Into Film)

This report is designed to account for progress during the third year of a four-year programme. Its purpose is to document delivery, report evidence against intended outcomes and show the extent to which the programme remained on course, rather than to provide a final account of long-term impact or definitive recommendations for the future.

That purpose shapes how the five building blocks appear:

- **Scene Setting** explains the programme's intended pathway from teacher development to film leadership, wider staff adoption and whole school change. It also establishes the continuing impact of COVID-19 on schools' capacity to deliver and embed the programme. *See Executive summary, About Film for Learning, About the report, and Section 1: Overview of the third year*
- **Evidence** and **Assessment** carry most weight. The report documents lesson delivery, teacher confidence, film leadership activity, reported pupil outcomes, school-level adoption and engagement with Into Film's wider services. It uses structured outcome-level judgements to indicate the extent of progress at this stage. *See Executive summary and Sections 1 to 4*
- **Significance** is present as an emerging account of whether the programme's sustainability pathway is beginning to work and why this is important. For example, teacher confidence matters because teachers were expected to lead and spread practice after formal support ended. *See the interpretation*

within Sections 2 to 4 and the future observations in Section 5: Concluding discussion

- **Implications** are appropriately provisional. The report identifies the questions to be resolved in the final evaluation, including whether practice is sustained, whether whole school change becomes embedded and whether early pupil and teacher benefits endure. *See Executive summary and Section 5: Concluding discussion*

Moving Minds 2: Supporting Young People's Mental Wellbeing (June 2021, Into Film)

This final evaluation report on Moving Minds 2 is intended to account for the delivery and outcomes of the completed project and to generate learning for future projects using a similar approach.

Its structure gives readers the detailed evidence before drawing together an overall assessment and implications for future practice.

- **Scene Setting** explains the challenge addressed, the project model, the intended beneficiaries and the varied delivery context. *See Executive summary, Introduction, and Outputs & Delivery*
- **Evidence** is given substantial weight. Monitoring data, feedback from participants and practitioners, and case study material describe what was delivered and how different people experienced it. This allows the report to show both positive

outcomes and the delivery conditions that shaped them.

See Outcomes and Appendix 1.

- **Assessment** follows from this evidence, reaching an overall view about the project's contribution while acknowledging variation between participants and settings. *See Executive summary, and outcomes-level commentary in Outcomes, and Discussion*
- **Significance** then identifies what the findings reveal about the approach: creative activity appears most effective when supported by trusted relationships, skilled facilitation, meaningful participation and organisational backing. *See Discussion*
- As a final report intended to inform similar future work, **Implications** are directed towards transferability. They identify which features of the model should be retained, what enabling conditions need to be in place, and what future projects should plan for from the outset. *See Discussion*

Judgement-led report examples

[BFI London Film Festival: Final Evaluation of the LFF 2020+ Strategy \(July 2024, BFI\)](#)

This is a final evaluation report on the BFI's LFF 2020+ strategy. It is designed to provide accountability for the Festival's performance against its strategic ambitions and to support learning for future editions.

The report therefore provides its principal judgements early, before presenting the supporting evidence and the implications for future planning.

The report is structured using a recurring pattern: establish relevant scene setting, state concise evaluative judgements, explain why each judgement matters, then supply selected evidence and use it to develop recommendations.

The executive summary and the "Summary assessment / Why does this matter? / What is the evidence?" architecture used in later chapters make this logic explicit.

- **Scene Setting** explains the Festival's cultural, audience, industry and talent development purposes, alongside the disruption and adaptation required during the pandemic. *See Executive summary, Introduction, Overview of LFF 2022 and 2023, and the introduction to sections 3, 4 and 5*
- **Assessment** is brought forward through headline conclusions about performance, contribution and strategic progress. *See Executive summary, headline findings and summary assessments in Sections 2, 3, 4 and 5, and Section 6: Value for money*
- **Significance** explains why the findings matter beyond event delivery: the Festival's contribution to cultural value, access to distinctive cinema, talent visibility and industry networking. *See 'Why does this matter?' in sections 2, 3, 4 and 5, and 6*
- **Evidence** then substantiates these judgements through audience data, stakeholder research, participation information

and qualitative accounts. *See 'What is the evidence?' in Sections 2, 3, 4, 5, and 6*

- Because this is a final report intended partly for learning, **Implications** can be relatively well developed. The report uses the completed strategy period to inform future Festival decisions, including how to build on effective practice and refine audience development, access, partnerships and digital provision. *See Section 7: Discussion and recommendations*

Mid-term Evaluation of Ofcom's Three-year Media Literacy Strategy 2024-2027 (July 2026, Ofcom)

Whereas the LFF report uses the five building blocks to structure the assessment of a mature festival strategy nearing its conclusion, this report shows how the same approach can support interim judgements when outcomes are still emerging and the primary purpose of the report is learning and course correction.

As an interim evaluation of Ofcom's three-year media literacy strategy, the report's purpose is twofold: to account for progress at the midpoint of the strategy period and to support learning and adjustment during the remaining term of delivery.

The report therefore provides an early assessment of progress while retaining a forward-looking focus on options for what should happen next.

- **Scene Setting** establishes the complex media literacy landscape and Ofcom's role within a wider ecosystem of

delivery organisations and policy development. *See Ofcom Foreword, 1. Introduction and opening paragraphs of sections 2, 3 and 4*

- **Assessment** is placed early in the report narrative, but is necessarily qualified: it identifies progress to date while recognising that some intended outcomes require more time and depend on partner activity. *See Executive summary and 'Overall assessment' in sections 2, 3 and 4*
- **Evidence** is used to support this account, covering delivery, stakeholder engagement, Ofcom's convening role and emerging change. *See 'Supporting evidence' in sections 2, 3 and 4*
- **Implications** are especially important because the report is mid-term. Rather than offering final conclusions, they identify where effort should be focused in the remaining period of the strategy. *See 'What could be continued, refined or developed further' in sections 2, 3 and 4, and 5. Conclusion and recommendations*
- **Significance** focuses on what the findings reveal about Ofcom's role (e.g. strengthening the evidence base and convening organisations matters only if this helps the wider system work more effectively for people and communities). *See 5. Conclusion and recommendations*

Glossary

Data collection plan

Sets out what evidence will be gathered during an evaluation, from which sources, and by what methods (surveys, interviews, monitoring data, and so on).

Evaluation questions (key evaluation questions)

The specific questions an evaluation has been commissioned or undertaken to answer, agreed in advance and used to keep analysis focused on what matters.

Evaluation rubric

Pre-agreed standards that describe what different levels of performance look like for a given outcome (e.g. what counts as "highly effective" versus "partially effective") used to make judgements more consistent and transparent.

Evaluative thinking and reasoning

A disciplined habit of mind that involves posing questions, seeking evidence, examining assumptions and reflecting critically before reaching a judgement.

Logic model

A visual or written summary of how a project's activities are expected to lead to its intended outcomes, usually simpler and less detailed than a full Theory of Change.

Monitoring data

Data collected routinely throughout a project's delivery (rather than specially for the evaluation), such as attendance figures.

Outcome

A change brought about by a project (e.g. improved confidence or better school attendance). This is distinct from an output (see below) and impact (the overall difference a project makes).

Output

Something tangible that directly results from a project (usually expressed as a measurable unit of delivery, e.g. number of sessions delivered, number of participants etc.)

Project logic / causal pathway

The assumed chain of cause and effect connecting a project's activities to its intended outcomes. It provides the reasoning for why a given action is expected to lead to a particular change.

Stakeholder

Anyone with an interest in, or affected by, the project or its evaluation (commonly including funders and commissioners, delivery staff, participants, and partner organisations).

Theory of Change

A detailed explanation of how and why a project is expected to bring about its intended outcomes, usually including the underlying assumptions and the conditions needed for success.

Triangulation

Comparing evidence from different sources or methods to see whether they corroborate or contradict each other, as a way of testing how robust a finding is.

Further reading

Evaluative thinking & reasoning

Evaluative Thinking: What is it? Why is it complex? Can We Boil it Down a Bit?,
Julian King, 2023

Evaluative Thinking,
NSW Government

Theory-based evaluation

Theory of Change in Ten Easy Steps,
NPC, 2019

Theory of Change,
NCVO, 2022

Theory-based Designs,
Evaluation Task Force Academy,
Cabinet Office

Theory-based Approaches to Evaluation: Concepts and Practices,
Government of Canada, 2009

Evaluation planning

Planning Your Evaluation,
NCVO, 2023

Rainbow Framework,
Global Evaluation Initiative

Planning and Managing Evaluations,
Evaluation Task Force Academy 2.0,
Cabinet Office

Evaluation Toolkit,
Ofcom, 2026

Evaluation Design Checklist,
Daniel L. Stufflebeam, The Evaluation
Center, Western Michigan University,
2004

Checklist of Key Considerations for Development of Program Logic Models,
Goldie MacDonald, The Evaluation
Center, Western Michigan University,
2018

Choosing a methodology

Choosing Your Evaluation Methods,
NCVO, 2023

Choose Methods and Processes,
Global Evaluation Initiative

Introduction to Mixed Methods in Impact Evaluation
Michael Bamberger, InterAction, 2012

Report writing

Checklist for Program Evaluation Report Content,
Kelly N. Robertson and Lori A. Wingate,
The Evaluation Center, Western Michigan
University, 2017

Checklist for Evaluation Recommendations,
Osvaldo Feinstein, The Evaluation Center,
Western Michigan University, 2019

Annex: Documentation for fictional projects used as examples in this guide

The following are simplified examples of the types of documentation used for planning a project and its evaluation.

They include just enough detail to support the example projects used in this guide to illustrate key principles and techniques associated with the five building blocks approach.

1. Brighter Start Eastmere (fictional reading support project)

Theory of Change

Problem the project seeks to address:

Nearly a third of children in Eastmere read below their expected age level by age seven, and the gap widens rather than narrows through primary school.

What the project plans to do about it:

Trained volunteers deliver one-to-one reading support, two hours a week, over two academic years, to children identified by teachers as reading below the expected level, across six primary schools.

The logic behind the project (causal pathway):

- IF children receive regular, individualised reading support → THEN their reading skills improve at a faster rate than they would through classroom teaching alone (because one-to-one attention allows pacing and content matched to the individual child, which whole class teaching cannot provide).
- IF reading skills improve → THEN reading confidence improves, reducing avoidance behaviours that further widen the gap.

- IF the gap narrows during the two-year window → THEN children are less likely to fall permanently behind as reading demands increase in later primary years.

What changes the project is expected to bring about:

INPUTS (what goes into the project)	OUTPUTS (tangible things the project delivers)	IMMEDIATE OUTCOMES (During project)	INTERMEDIATE OUTCOMES (At the end of 2 years)	LONGER-TERM OUTCOMES (1-3 years after project)	WIDER IMPACT
• Funding • Volunteer readers • Volunteer supervision capacity • Teacher time	• Number of children enrolled across the six schools • Number of volunteers trained and actively delivering sessions • Number of one-to-one sessions delivered per child	1. Improved reading accuracy and fluency 2. Increased reading confidence	3. Reduced gap between participating children's reading age and chronological age 4. Reduced avoidance behaviours around reading in the classroom	5. Sustained reading-age gains maintained into later primary years 6. Improved engagement with reading-dependent subjects across the curriculum	Narrower reading attainment gap across Eastmere's primary schools Less need for more intensive literacy interventions at secondary level Stronger case for sustained investment in early literacy support as a preventative measure, rather than later remediation

Assumptions about how the project will work:

1. The main barrier to reading progress for this group is insufficient individualised attention, not, for example, undiagnosed learning needs, home environment or English as an additional language.
2. Regular one-to-one reading support, delivered consistently over time, could close the reading age gap before it hardened into a longer-term disadvantage.
3. Volunteers, with training but not professional teaching qualifications, can deliver support of sufficient quality to produce measurable gains.
4. Two hours a week is sufficient contact time to produce a meaningful effect within two years.
5. Consistent attendance is achievable once a child is enrolled.
6. Gains made during the project will be sustained after it has finished, rather than fading as soon as one-to-one support ends.

Key Evaluation Questions

1. To what extent did participating children achieve greater reading age gains than would be expected without the project?
2. Did the size of that effect vary according to attendance levels, and if so, how?
3. What did teachers and children observe about changes in reading confidence, beyond what reading age assessments capture?
4. Which of the project's underlying assumptions were supported or challenged by the evidence?
5. Were gains sustained in the term following the end of one-to-one support?

Data Collection Plan

Method	What it captures	Linked to...
Reading age assessment (standardised tool)	Baseline, midpoint, and end-of-project reading age scores per child	Immediate/intermediate outcomes; the reading age rubric
Attendance register	Sessions attended vs. scheduled, per child, per volunteer	Output tracking; subgroup analysis by attendance level
Teacher short survey (2 points: midpoint, end)	Observed changes in reading confidence and classroom engagement	Immediate outcomes not captured by reading age scores alone
Volunteer session log	Brief notes per session: content covered, engagement level, any concerns	Fidelity of delivery; early flag for attendance or engagement issues
Volunteer feedback survey	Volunteer perspective on what worked, delivery challenges, training adequacy (end of project)	Process evaluation
Enrolment records	Number of children referred, eligible, and enrolled per school	Output tracking; equity of access across the six schools

Evaluation rubric for intermediate outcome 3:

'Reduced gap between participating children's reading age and chronological age'

Standard	Threshold
Exceeded	Average gain across participants of 8+ months of reading age over the two-year project, with no identifiable subgroup showing negligible gain.
Met	Average gain of 5–7 months, broadly consistent across attending participants (allowing for normal individual variation).
Partially met	Average gain of 5+ months overall, but with a clearly identifiable subgroup showing negligible or no gain.
Not met	Average gain at or below the expected gain for non-participating peers (4 months), or no reliable evidence of an effect.

2. Step Forward Now (fictional employment support project)

Theory of Change

Problem the project seeks to address:

Adults unemployed for 12+ months in Eastmere face compounding barriers to re-entering work, and standard job centre support has limited success with this group.

What the project plans to do about it:

A six-month project combining weekly one-to-one coaching with access to short, funded vocational courses (initially in care work and logistics), for adults referred by local job centres after 12+ months of unemployment.

The logic behind the project (causal pathway):

- IF participants gain relevant vocational skills through funded courses → THEN they become more competitive candidates for available roles in Eastmere's growth sectors (care, logistics).
- IF participants receive regular coaching → THEN their confidence, motivation and job search persistence improve, counteracting the demotivating effects of long-term unemployment.
- IF skills and confidence improve together → THEN participants are more likely to secure and sustain new work within six months of completing the project.

What changes the project is expected to bring about:

INPUTS (what goes into the project)	OUTPUTS (tangible things the project delivers)	IMMEDIATE OUTCOMES (During project)	INTERMEDIATE OUTCOMES (At the end of 2 years)	LONGER-TERM OUTCOMES (1-3 years after project)	WIDER IMPACT
- Funding - Recruited and trained employment coaches - Referral partnership with local job centres	- Number of participants referred and enrolled per cohort - Number of coaching sessions delivered per participant - Number of participants completing funded vocational courses	- Increased participant confidence and motivation - Acquisition of vocational skills - Increased job search activity	- Increased rate of job offers and job entry relative to national benchmark - Improved participant self-reported readiness for sustained employment - Reduced reliance on job centre support among participants who secure work	- Sustained employment beyond initial job entry - Reduced likelihood of returning to long-term unemployment - Increased household income stability for participants and their families	Reduced long-term unemployment rate in the local area, particularly among the 12+ months unemployed cohort Reduced demand on local welfare and job centre support services over time Stronger local evidence base for whether coaching-led or skills-led models are more effective for long-term unemployed adults, informing future commissioning decisions

Assumptions about how the project will work:

1. The primary barrier to employment for this group is a skills deficit relative to available roles. *(This is the assumption that positions vocational courses as the project's central lever).*
2. Coaching functions mainly as a supporting mechanism to sustain engagement with the courses and job search, rather than as an independent driver of outcomes.
3. Six months is sufficient time for a participant to complete a course, build confidence, and secure work.
4. Growth in care and logistics sector vacancies locally is sufficient to absorb newly skilled participants.
5. Length of prior unemployment does not meaningfully change how participants respond to the project. *(i.e. The model assumes a fairly uniform effect regardless of how long someone has been out of work, although this was flagged as untested at design stage).*

Key Evaluation Questions

1. What proportion of participants entered work within six months of completing the project, and how does this compare to a relevant benchmark?
2. Does length of prior unemployment affect the likelihood of a positive outcome?
3. What do participants and coaches identify as the most valuable components of the project?
4. To what extent is the assumed causal role of vocational courses (as the primary driver of employment outcomes) supported by the evidence, relative to coaching?
5. Are there early indications of sustained employment beyond the six-month mark? *(Also flagged in the original design as a question the six-month evaluation window could only partially answer.)*

Data collection plan

Method	What it captures	Linked to...
Case management system extract	Coaching sessions attended, course enrolment/completion, job outcomes at 6 months	Output tracking; job entry rubric
Job entry and retention check (6 months post-project)	Employment status, hours, and job type at the 6-month mark	Intermediate outcome; core evaluation question on job entry
Coach case notes (ongoing, reviewed at project end)	Qualitative record of participant progress, confidence and barriers raised in coaching	Evidence on coaching's role; Process evaluation
Participant exit survey	Satisfaction with coaching and courses; self-reported confidence and job search readiness	Evaluation question on valued components
Course completion records (provider data)	Completion rates and any course level outcomes (e.g. certification)	Output tracking
Referral and enrolment records (job centre data)	Length of prior unemployment per participant at referral	Subgroup analysis by unemployment duration

Evaluation rubric for intermediate outcome 4:

'Increased rate of job offers and job entry relative to national benchmark'

Standard	Threshold
Exceeded	Job entry rate at 6 months exceeds national benchmark ($\approx 20\%$) by 15 percentage points or more, with no significant disparity by prior unemployment duration.
Met	Job entry rate exceeds benchmark by 5–14 percentage points, with only modest variation across subgroups.
Partially met	Job entry rate exceeds benchmark overall, but with a sizeable minority (one third to one quarter) of participants showing materially weaker outcomes.
Not met	Job entry rate at or below national benchmark, or no reliable evidence of a project effect over what job centre support alone would achieve.

2006 - 2026
Celebrating 20 years
www.biggerpictureresearch.com

