verian

JK&A

A Guide to Evaluation of Value for Money in UK Public Services

Why cost-benefit analysis alone may be insufficient to evaluate VFM, and how to navigate a solution

Julian King & Alex Hurrell

August 2024

Contents

Foreword

At Verian, renowned for our research, evaluation and specialised consultancy services, we support the use of cost-benefit analysis (CBA) as a crucial tool in our approaches to delivering data-driven analysis. In combination with our insights into human behaviour, attitudes and experiences, the team at Verian leverages CBA to enhance policy making and social impact. However, CBA is not a panacea. We recognise its value while asserting that CBA alone isn't always sufficient.

As we grapple with the complexities of assessing value for money (VfM), we find that the possibility of merging insights from CBA with other methods and tools from programme evaluation offers a compelling avenue for gaining a nuanced understanding of policies and programmes and supporting sound decision-making. In this Guide, we introduce *Value for Investment*, an approach that has been gaining traction globally and has recently been integrated into UK public sector evaluation training by the HM Treasury/Cabinet Office Evaluation Task Force.

In this document we challenge a conventional notion that equates VfM with CBA. We argue that VfM can be evaluated in various ways, including but not limited to CBA and that methods should be selected according to context. While acknowledging the importance of economic analysis, this Guide advocates for a more inclusive approach that considers a spectrum of values – balancing return on investment with other considerations like equity and affordability – alongside a diverse range of evidence sources, both quantitative and qualitative. By applying this multifaceted VfM evaluation methodology, we can enhance the depth and validity of VfM assessment without confining it to economic metrics alone.

In a landscape where VfM is receiving increasing scrutiny, and evidence-based decision-making is paramount, we hope that this Guide will serve as a navigational aid for those who are looking to design and deliver comprehensive and context-appropriate VfM assessments. We hope it will inform and inspire a new perspective in the use of sound evaluation practices to assess VfM, embracing diversity in perspectives and richness in evidence sources for informed and impactful policy outcomes can maximise the future public value of all government spending.

Craig Watkins
UK Chief Executive Officer, Verian, London

Executive Summary

When assessing value for money (VfM), there are sound arguments for combining insights from economics and evaluation. Drawing on the complementary strengths of both disciplines supports comprehensive assessment of policies and programmes' VfM. In this Guide we introduce an inter-disciplinary approach, *Value for Investment*, that integrates methods and tools from both disciplines, and which has gained traction worldwide.

Among the concepts and methods of evaluation that economics brings to this interdisciplinary mix is cost-benefit analysis (CBA), which evaluates whether an intervention creates more value for society than it consumes. CBA is vital in the UK policy context, acting as a key tool for evaluating public policies, programmes, and projects as set out in the *Green Book,* HM Government's central guide for appraisal and evaluation, and the *Magenta Book* which focuses on evaluation methodologies and practices across the whole policy cycle. Together, the Green Book and Magenta Book form a robust framework for evidence-based policymaking, ensuring CBA informs sound decision-making and evaluates implemented policies' impacts.

Within this policy context, value for money (VfM) is often framed as an objective or question within an evaluation and is commonly equated with CBA. However, we argue that while CBA contributes important evidence, it may be insufficient. The approach set out in this Guide includes economic analysis while integrating multiple values (social, economic, environmental, and cultural) and diverse evidence sources (quantitative and qualitative) for a more comprehensive VfM assessment, drawing on the strengths of CBA without limiting the evaluation to economic methods and metrics alone.

The Value for Investment approach uses mixed methods (integrating quantitative and qualitative evidence), employs evaluative reasoning (interpreting evidence through the lens of explicit criteria and standards – improving transparency in the rationale for evaluative judgements), and is participatory (involving stakeholders in co-design and analysis). It is designed to be intuitive and practical to use by following a logical sequence of steps.

The Value for Investment approach was developed through doctoral research (King, 2017; 2019) and refined through real-world VfM evaluations in diverse contexts. Leading examples of its application are published in *Assessing Value for Money – the Oxford Policy Management Approach* (King, Wate, Namukasa, Hurrell, Hansford, Ward, & Faramarzifar, 2023), and *Value for Investment – Application and Insights* (King, Crocket, & Field, 2023).

This Guide is designed to complement existing Vfl guidance by focusing on the interface between CBA and other evaluation methods to provide sound answers to VfM questions in the context of UK public policy.

1. Background

When evaluating value for money (VfM), it is beneficial to integrate perspectives from both economics and evaluation. Utilising methods and tools from both disciplines allows for a thorough assessment of policies and programmes. In this guide, we present an interdisciplinary approach called *Value for Investment*, which employs mixed methods (qualitative, quantitative, and economic) alongside evaluative reasoning (interpreting evidence through the lens of explicit criteria and standards) to deliver more nuanced and comprehensive VfM assessments.

1.1. UK Policy Context

Value for money (VfM) assessment is important for accountability and transparency in the use of public resources as well as for learning, improvement, and communicating the value of policies and programmes (King et al., 2023). There is no universal definition of VfM (King, 2019). However, in the UK, the National Audit Office (NAO) has defined VfM as "the optimal use of resources to achieve the intended outcomes" (NAO, n.d.).

Cost-benefit analysis (CBA) plays a crucial role in the UK policy context, serving as a fundamental tool for evaluating and appraising public policies, programmes, and projects. The *Green Book* (HM Treasury, 2022a), which provides central government guidance on appraisal and evaluation, is the primary resource for conducting CBA in the UK. This comprehensive guide outlines the methodology for assessing the costs, benefits, and trade-offs of various policy options, ensuring that public resources are allocated efficiently and effectively. The Green Book emphasises the importance of considering both monetary and non-monetary impacts, including environmental and social factors, to provide a holistic view of a policy's potential outcomes.

Complementing the Green Book is the *Magenta Book* (HM Treasury, 2022b), which focuses on evaluation methodologies and practices. While the Green Book provides guidance on *ex-ante* (before the event) appraisal, the Magenta Book offers insights into how to design, conduct, and interpret evaluations throughout the policy cycle with the aim of assessing whether expected impacts and benefits have materialised. Together, these two resources form a robust framework for evidence-based policymaking in the UK, ensuring that CBA is not only used to inform initial decision-making, but also to assess the actual impacts of implemented policies. This comprehensive approach allows policymakers to refine and improve interventions based on empirical evidence, ultimately leading to more effective and efficient public policies.

In the UK public policy context, VfM is often framed as an evaluation objective or question within a wider evaluation (King, 2023i). For example, terms of reference might call for an evaluation of processes, outcomes, and VfM. In this setting, the term 'value for money' is often used almost interchangeably with economic evaluation methods, such as CBA. However, VfM and CBA are not synonymous. VfM can be assessed in various ways including but not limited to CBA (Gargani & King, 2023). We argue that in many public investments it would often be too limiting to rely on CBA alone (King, 2017). For example, CBA assesses performance against a single criterion (the monetised value of impacts relative to costs) whereas VfM assessment often balances multiple criteria that can relate to different stages of a policy or programme. Moreover, the scope of a CBA may be constrained by what impacts can be measured and monetised. While the output of a CBA can provide valuable evidence, it may not be sufficient on its own to provide a comprehensive VfM assessment. We argue that to effectively evaluate VfM, the UK public sector should utilise a broader range of evaluation tools in conjunction with CBA.

The NAO and the Foreign, Commonwealth & Development Office (FCDO) appear to agree, recommending a combination of up to five criteria to assess VfM of government spending: economy

(minimising input costs while having regard to quality), efficiency (the relationship between outputs and the resources used to produce them), effectiveness (achieving objectives and outcomes), equity (the extent to which services benefit the intended people), and cost-effectiveness (the relationship between resource use and impacts) (DFID, 2011; NAO, 2024). While the 'cost-effectiveness' criterion may be addressed by using CBA, this is not a requirement, nor is it always feasible. Moreover, the inclusion of five criteria necessitates trade-offs and judgements, making VfM a matter of context and perspective (King et al., 2023).

1.2. Purpose of this Guide

In this Guide, we outline an approach called *Value for Investment* (VfI) (King, 2017; 2019), which includes economic analysis but goes beyond it, offering a framework that integrates multiple values (e.g., social, economic, environmental, and cultural) and diverse evidence sources (quantitative and qualitative) for a more comprehensive perspective on the costs, processes, outcomes, and value of policies and programmes.

This Guide is designed to complement existing VfI guidance including *Assessing Value for Money – the Oxford Policy Management approach*[1] (King et al., 2023) and *Value for Investment – application and insights*[2] (King, Crocket, & Field, 2023).

The unique contribution of this Guide is its focus on the interface between CBA and the full suite of evaluation methods in the context of UK public policy, including:

- How CBA can make a valuable contribution to VfM assessment;
- Why CBA alone may not be enough;
- How to conduct a VfI evaluation that matches methods to context and includes CBA where feasible and appropriate; and
- How to use VfI as a viable alternative when CBA isn't possible.

[1] King, J., Wate, D., Namukasa, E., Hurrell, A., Hansford, F., Ward, P., Faramarzifar, S. (2023). *Assessing Value for Money: the Oxford Policy Management Approach*. Second Edition. Oxford Policy Management Ltd.

[2] King, J., Crocket, A., Field, A. (2023). *Value for Investment: Application and Insights*. Youth Primary Mental Health and Addictions Evaluation. Exemplar report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Ltd.

2. Why not just do a cost-benefit analysis?

Cost-benefit analysis (CBA) offers unique insights that are not available through other means and is a core evaluation method. However, CBA alone may not offer a comprehensive evaluation of policies and programmes where outcomes are complex and hard to monetise. Integrating CBA with mixed methods can enhance both approaches, leveraging their respective strengths to provide more robust evaluations. This section explains why.

2.1. What is CBA?

CBA is an economic method of evaluation, used to compare the potential or actual gains and losses to society attributed to a policy, programme or initiative. It involves systematically identifying and valuing benefits and costs in monetary terms. The values of future benefits and costs are converted to present values using a *discount rate* which reflects the premium a society places on the present over the future (HM Treasury, 2022).

CBA is commonly used in financial decision-making, where records or forecasts of capital investments, operating costs, and generated revenues, along with scenario analysis, support investment and operational decisions based on estimated risk and return on investment. This approach is sometimes referred to as 'financial CBA'. However, CBA is more than just a financial analysis tool. It is widely used in policymaking across various sectors such as transport, infrastructure, healthcare, education, and community services. When CBA is conducted to include intangible social values, such as social benefits and costs, it is often termed 'social CBA' (King, 2023p).

Conducting a CBA involves:

- Identifying incremental costs and benefits attributable to the intervention, programme, or policy, relative to a counterfactual;
- Quantifying and valuing them in monetary terms (drawing values from a range of sources such as financial accounts, market prices, choice experiments, willingness-to-pay or subjective wellbeing surveys, literature, and values banks) (King, 2024c; 2024d);
- Arranging them in a time series, according to their occurrence over an appropriate time horizon;
- Weighting the costs and benefits to adjust their values for timing (i.e., *discounting*), and sometimes adjusting for other factors like risk; and
- Aggregating the discounted costs and benefits to produce indicators such as *net present value* (NPV) or *benefit-cost ratio* (BCR).

The NPV is equal to the present value of benefits minus the present value of costs. The BCR is equal to the present value of benefits divided by the present value of costs. If NPV is greater than zero, or BCR is greater than one, it may be concluded that the investment is worthwhile on the basis that it creates more value than it consumes.

However, the analyst may take additional factors into account when judging whether an investment is worthwhile – such as the degree of uncertainty and/or risk involved, assessment of distributional impacts, and any significant intangible benefits or costs that weren't included in the analysis. These factors may be weighted implicitly or explicitly. We recommend that the reasoning behind the judgements should be explicit, and that the inclusion of multiple methods and considerations should be the norm. The next chapter expands on these recommendations.

CBA can be used to look forward (*ex-ante*) to inform business cases and decisions about future investments (as detailed in the *Green Book*), or backward (*ex-post*) to assess whether benefits and costs turned out as expected (as described in the *Magenta Book*). These guides rightly position CBA as an important method in VfM assessment.

A robust and credible CBA can be data-intensive, requiring a defined counterfactual, measures for relevant outcomes, data on intervention costs and outcomes, data or estimates of counterfactual costs and outcomes, and valuation factors to translate outcomes into monetary values (Schiff, 2024). Where a CBA (or any other evaluation) is anticipated, we recommend building evaluation planning into programme design to ensure the necessary data and information are available at the right time.

2.2. Strengths of CBA

Cost-benefit analysis (CBA) holds a high status in policy making, reflecting its advantages as an evaluation method. Strengths of CBA include:

- **Unified valuation**: CBA values costs and benefits in monetary units, allowing direct reconciliation into a single indicator such as NPV or BCR.
- **Compelling communication**: NPV and BCR enable value to be communicated in a simple way that is often compelling for stakeholders and decision-makers (for example, a BCR may be communicated as: "for every £1 invested, £3 of social value is created").
- **Versatility**: CBA can be used to assess and compare costs and benefits of almost any public investment in any sector or setting, including both *ex-ante* appraisals and *ex-post* evaluations.
- **Timing adjustment**: CBA adjusts for timing by discounting costs and benefits, reflecting societal preferences for present versus future value.
- **Sensitivity analysis**: CBA incorporates sensitivity and scenario analyses, providing insights into the impact of varying assumptions and uncertainties on evaluation results (see Section 7.1).
- **Consistency and transparency**: CBA adheres to well-established guiding principles that ensure consistency, transparency, and replicability across evaluations, fostering confidence in its findings.
- **Overall societal benefit**: CBA addresses an important question: whether society is better-off overall from a given intervention (King, 2023e).

CBA has long been the preferred method for assessing the economic efficiency and viability of policies, programmes, and interventions, to such an extent that VfM is sometimes mistakenly equated with CBA. However, VfM and CBA are not synonymous, and conflating them can impede effective resource allocation decisions. VfM can be evaluated using a variety of methods, including but not limited to CBA (Gargani & King, 2023). Despite CBA's strengths, we argue that evaluating complex public investments often requires supplementary methods alongside CBA to provide more comprehensive VfM assessment and decision-making support.

2.3. Limitations of CBA

The use of CBA in policy and programme evaluation entails complexities and issues that warrant careful consideration. Key examples of limitations and challenges in the use of CBA include:

- **One guiding criterion**: Generally in evaluation, criteria are determined according to context and multiple criteria are the norm (Schwandt, 2015). In CBA, there is one criterion, which is an

integral part of the method and is not always made explicit. CBA is primarily concerned with whether an intervention makes society better off in the aggregate. Any change that produces total benefits greater than total costs is considered to be worthwhile, regardless of how the costs and benefits are distributed across society. This criterion is formally known as *Kaldor-Hicks efficiency.* Additional, non-efficiency perspectives (such as equity considerations, cultural relevance, sustainability, and broader societal impacts) are essential for comprehensive and fair resource allocation decisions (King, 2017). While CBA can include distributional assessment (HM Treasury, 2002a) – for example, by including subgroup analyses and equity weights – these elaborations to CBA can only supplement and should not replace more comprehensive analysis of equity-related issues (King, 2023h).

- **Challenges in valuing intangible benefits and costs**: Social CBAs often exclude intangible values when they are too difficult to measure in monetary units (Adler & Posner, 2006). For example, benefits like community cohesion or cultural heritage preservation, which are valuable yet challenging to monetise, may be excluded. Even when monetary valuation is feasible in principle, some intangible values may be left out of the analysis in order to achieve a reasonable degree of precision and credibility (King, 2024c). In social investments, some of the most valuable benefits may be among the hardest to monetise. Consequently, CBAs may provide an incomplete assessment of some social programmes and policies.

- **Aggregation of values**: CBA consolidates diverse values into a standard metric, allowing multiple different factors like health, wealth, and happiness to be considered within a unified framework. However, doing so can oversimplify complex situations and runs the risk of overlooking qualitative nuances. For example, aggregating values isn't always appropriate, potentially masking underlying tensions and power disparities in decision-making processes (Julnes, 2012). Unless used in combination with other methods, aggregating values bypasses nuanced deliberation and human judgement essential in addressing complex social issues. For example, public dialogue and collective values are peripheral to CBA (King, 2017; 2019). Moreover, assigning monetary valuations to benefits and costs renders them *fungible* – able to be traded off against anything else with a monetary valuation. However, some values (such as ethical bottom-lines) should not be included in such trade-offs (King, 2023f).

- **Consequentialist perspective**: CBA is conducted from a consequentialist viewpoint, focusing on the outcomes and impacts of decisions relative to their costs while ignoring the organisational actions and outputs that conceptually link resources to outcomes in a theory of change. This emphasis makes CBA potentially unsuitable for assessing VfM in the early stages of a policy or programme and may undervalue the importance of procedural fairness and the intrinsic value of processes themselves in decision-making (Julnes, 2012).

- **Responsiveness to stakeholders:** CBA can be, and often is, conducted on a desktop basis, without engaging stakeholders to understand what they value, nor to seek their input into what impacts may materialise (and how), and what methods are appropriate and acceptable. CBA *can* be conducted in a manner that is inclusive, responsive, contextually viable and meaningful for stakeholders; there's just nothing in the CBA methodology that says it *should.* CBA, like any evaluation method, should be conducted in keeping with programme evaluation standards (e.g., UKES, n.d.; Yarbrough et al., 2011) – and that requires remaining open to the possibility of using it in combination with other methods or, in some cases, *not* doing a CBA (King, 2023a, 2023b).

- **Risk of unwarranted authority:** A compelling statement like "for every £1 invested, £3 of social value is created" can carry a lot of weight in decision-making and in justifying and communicating decisions. There is an onus on evaluators and decision-makers to ensure that

any such emphasis is warranted and to guard against inappropriate use, bearing in mind the above-mentioned limitations.

In summary, CBA has multiple limitations and therefore is not always the appropriate method for assessing VfM. While CBA contributes an important metric to inform decision-making, it is essential to recognise its limitations in capturing the full spectrum of values and considerations in complex social evaluations. These are not arguments against using CBA – but they do suggest that in at least some circumstances it is desirable to combine CBA with other methods, harnessing its strengths while compensating for its limitations (King, 2017). Combining multiple evaluation methods within a broader framework can inform more holistic resource allocation decisions (King, 2019).

2.4. A broader approach is necessary

Some economists and public policy experts have suggested that insights from CBA should be integrated with broader considerations to guide decision-making. These arguments recognise both the strengths and limitations of CBA. For example, Adler & Posner (2006) described CBA as an imperfect but useful decision procedure. Sunstein, in *The Cost-Benefit Revolution* (2018, p. xi) took a similar view, arguing that CBA provides essential information, even if that information is incomplete:

> *"Whether or not an analysis of costs and benefits tells us everything we need to know, at least it tells us a great deal that we need to know. We cannot safely proceed without that knowledge."*

These perspectives also recognise that essential information for policymaking extends beyond technocratic approaches to analysis and includes democratic processes such as citizen engagement (Kahneman, 2011). Building on this perspective, Flyvbjerg and Bester (2021), in their paper *The Cost-Benefit Fallacy*, stated:

> *"For cost-benefit analysis to be accepted and have impact, it must be understood… in ways that fit with the messy, non-expert character of present-day democratic decision making. Here cost-benefit analysis is just one of many inputs that are amalgamated into the overall decision-making process."*

Similarly, Adler and Posner (2011), in *New Foundations for Cost-Benefit Analysis*, recognised that there are values beyond those captured by CBA. They concluded that "CBA is not a superprocedure" for combining all relevant values and acknowledged the need for a comprehensive approach to integrate CBA outputs with wider considerations. They admitted:

> *"We suppose that that is a theoretical possibility—but we have absolutely no idea what the superprocedure would consist in."*

The field of programme evaluation offers methods and tools to address this need. The superprocedure that eluded Adler and Posner is known as *explicit evaluative reasoning*, which is central to the practice of evaluation (King, 2017).

Evaluative reasoning is the process by which evaluators reach a warranted value judgement – that is, a transparent judgement supported by evidence and logical argument (Davidson, 2005). It is generally considered a good practice to make reasoning explicit in evaluation (Yarbrough et al., 2011). In the next section, we unpack how to design and implement evaluations that are underpinned by explicit evaluative reasoning.

3. The Value for Investment approach

This section explains how to design and implement VfM assessments that can include CBA within a broader mix of methods. Value for Investment (VfI) isn't a method – it's an evaluation system developed to bring clarity to using existing methods and tools to evaluate how well resources are used, the extent to which value created justifies the resources invested, and how more value can be generated from investments in policies or programmes.

Developed through doctoral research (King, 2017; 2019) and refined through real-world VfM evaluations in diverse contexts (e.g., King & Guimaraes, 2016; King & Allan, 2018; King & OPM, 2018; King et al., 2023; King, Crocket, & Field, 2023; Kinnect Group & Foundation North, 2016; Oakden & King, 2018), VfI is in use globally to evaluate complex and hard-to-measure policies and programmes.

The inclusion of the VfI approach in evaluation training by the UK's HMT/Cabinet Office Evaluation Task Force, and recognition through awards by the Australian and UK Evaluation Societies, underscore its innovative nature and effectiveness in assessing VfM.

3.1. Guiding principles

The VfI system is underpinned by four key principles: it is interdisciplinary (drawing on evaluation and economics), uses mixed methods (integrating quantitative and qualitative evidence), employs evaluative reasoning (interpreting evidence through explicit criteria and standards), and is participatory (involving stakeholders in co-design and analysis) (King, 2019).

3.1.1. Inter-disciplinary

Value for Investment (VfI) is an interdisciplinary approach, combining theory and practice from evaluation and economics to determine whether a policy or programme provides VfM. Evaluation provides systematic assessments of merit and worth – determining how ***good*** something is and whether it is good enough (Davidson, 2005), while economics studies how people choose to use ***resources***. VfM (***good resource use***) sits at the intersection of the two disciplines (King, 2017; 2019).

Evaluation's contributions to the VfI approach include evaluative reasoning, mixed methods, and participatory approaches (as outlined in the sections below), together with "systems thinking and complexity-informed approaches, and a diverse toolbox for incorporating multiple values, forms of evidence, and approaches to causal inference" (King et al., 2023). From the field of economics, concepts such as opportunity cost, efficiency, productivity, cost-effectiveness analysis, and CBA (to name a few) contribute to VfI frameworks and methods.

By integrating insights from both disciplines, VfI leverages their complementary strengths, concepts, models, frameworks, and tools. For example, if there are considerations that defy monetary valuation, they don't have to be consigned to the footnotes of a CBA; the field of evaluation offers multiple methods and tools for working rigorously with evidence and values, accommodating more complexity and nuance than CBA alone. While economic methods like CBA are useful for assessing efficiency, a comprehensive VfI evaluation also considers equity and other criteria.

3.1.2. Mixed methods

Mixed methods, incorporating both quantitative and qualitative evidence, can enhance evaluations by providing a deeper understanding than single-method approaches. For example, mixed methods can: complement each other, allowing for triangulation of evidence to identify converging or diverging insights; leverage the strengths of diverse evidence sources to better understand value

creation processes, enhancing the reliability and validity of findings; and enable the design of one method to inform another (Greene, 2007).

Consequently, mixed methods facilitate a more comprehensive and nuanced understanding that conveys the story behind the numbers. Within a VfI evaluation, the mix of methods can include CBA or other economic methods of evaluation where feasible and appropriate. VfI also offers a viable alternative in circumstances where CBA is considered infeasible or unnecessary by drawing on a wider range of evaluation methods (King, Crocket & Field, 2023).

3.1.3. Evaluative reasoning

Evaluation is essentially determining the value (e.g., merit, worth, or significance) of something like a policy, programme, or intervention. It involves making value judgements, and those judgements should be transparent and *warranted*, supported by evidence and logical argument. *Evaluative reasoning* is the process by which evaluators reach a warranted judgement (Fournier, 1995).

In professional evaluation this process can be made transparent by combining empirical evidence with explicit criteria (aspects of value) and standards (levels of value). The evaluative reasoning process has four components: identifying criteria, developing standards, gathering and analysing evidence, and synthesising the evidence through the lens of the criteria and standards to make evaluative judgements (Fournier, 1995).

Criteria and standards may be set out in a *rubric* (Davidson, 2005; King, 2023n) articulating a transparent basis for making evaluative judgements, developed with stakeholders and agreed during the design phase of the evaluation (King, 2019). Used well, rubrics can support clarity and transparency at every step of an evaluation, from design to data gathering, analysis, synthesis and reporting (King et al., 2013).

3.1.4. Participatory

Criteria and standards are determined on the basis of values (Schwandt, 2015), and values are "an expression of what matters to people" (King, 2019). In other words, "things don't have value; people place value on things" (Gargani & King, 2023). It follows that careful consideration must be given to which people and groups should be involved in developing criteria and standards that will have face validity with stakeholders. Appropriate and proportionate processes for stakeholder involvement are determined contextually (King et al., 2023).

An inclusive, participatory approach to evaluation involves engaging stakeholders, end-users, and communities to ensure their voices are heard and their values are included. This contributes to the credibility and validity of the evaluation framework and findings, enhances stakeholders' understanding of the evaluation process, and increases the likelihood that the evaluation will be endorsed and used (King et al., 2013; Patton & Campbell-Patton, 2021). Participatory approaches ensure evaluations meet standards of inclusivity, responsiveness, and contextual relevance (King, 2023a; Yarbrough et al., 2011). Evaluations conducted collaboratively with stakeholders or led by stakeholders may be more likely to have positive consequences than evaluations done to or for stakeholders (Wehipeihana, 2019).

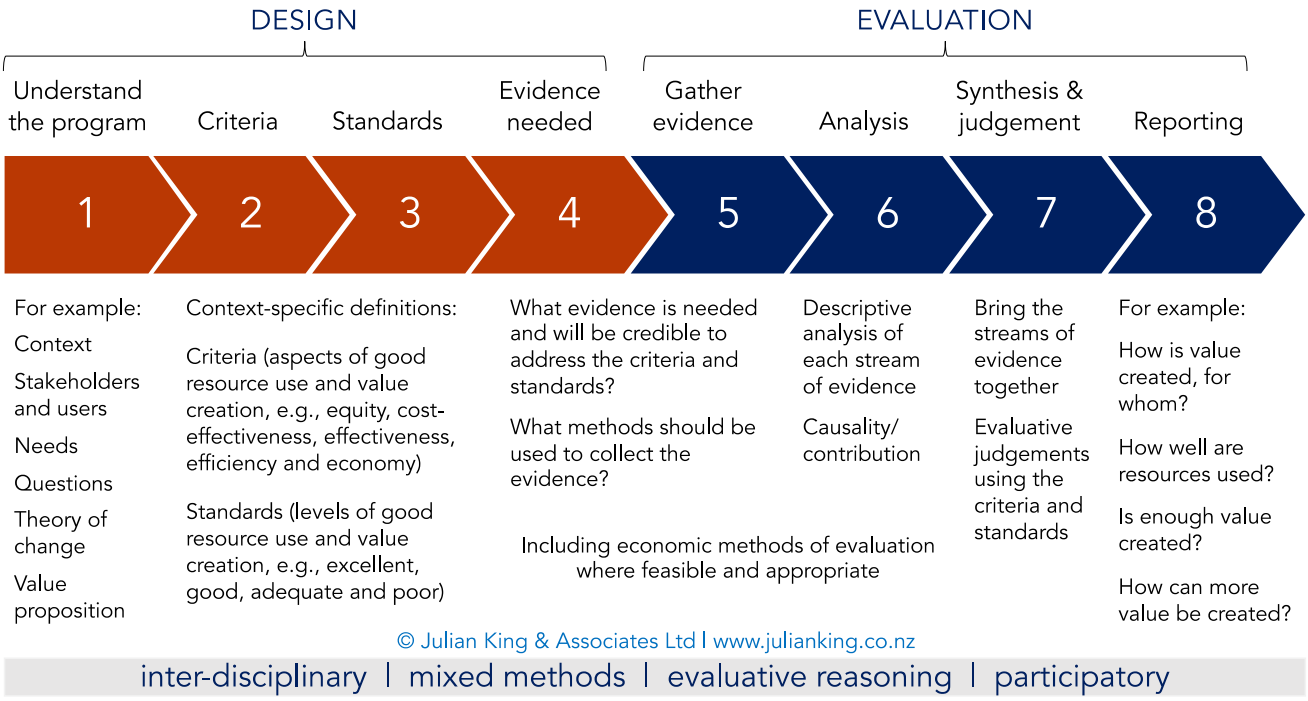
3.2. Practical process

The VfI evaluation process is designed to be intuitive to adopt and practical to implement, even for those new to evaluation, by following a logical eight-step sequence (King, 2019). This process incorporates the four components of evaluative reasoning – criteria, standards, evidence, and synthesis – within its eight steps.

The process is divided into two main phases: design and implementation. The design phase involves the four steps of: understanding the programme, defining its value proposition and evaluation questions (step 1); setting criteria for good resource use (step 2); establishing standards for these criteria (step 3); and determining the necessary evidence (step 4) – including economic methods of evaluation where feasible and appropriate.

The implementation phase comprises the subsequent four steps of: gathering evidence (step 5); analysing the evidence (step 6); interpreting the evidence using the defined criteria and standards (step 7); and presenting clear, evidence-supported answers to the evaluation questions (step 8). This structured approach ensures a comprehensive and accessible evaluation process.

Figure 1: The Value for Investment process



Detailed descriptions of the VfI process are well covered in existing resources such as *Assessing Value for Money – the Oxford Policy Management Approach* (King et al., 2023), and *Value for Investment – Application and Insights* (King, Crocket, & Field, 2023). The following summary focuses on determining whether and how to incorporate CBA within a VfI evaluation.

3.2.1. Evaluation design

The first four steps of the VfI process are the evaluation design steps of understanding the programme, developing criteria, setting standards, and identifying what evidence and methods are appropriate. These steps are generally carried out through participatory processes with stakeholders (King, 2023g).

Step 1 – Understand the programme

At the beginning of any evaluation, it's crucial to gain a thorough understanding of the programme, its context, key users, stakeholders, and their information needs (Scriven, 2013; Patton & Campbell-Patton, 2021). These information needs shape the evaluation questions (Davidson, 2005). This initial phase involves establishing a clear and mutual comprehension of the programme, including its intended goals and functionality.

Developing a theory of change helps to clarify the programme and ensures stakeholders' expectations and assumptions are aligned. A theory of change describes anticipated processes and outcomes (Funnell & Rogers, 2011), explaining how resources or inputs (like staff, policies, and knowledge) facilitate activities and outputs, which ultimately lead to outcomes and broader impacts for participants, communities, society, and government.

In a VfI evaluation, it can be helpful to expand a theory of change by specifying the programme's value proposition (King, 2021; King, Namukasa & Hurrell, 2023). This innovative addition to programme theory makes value creation explicit by addressing questions such as:

- To whom is the investment valuable, and in what ways?
- What resources are invested, by whom? Who needs the investment, and why?
- How does the investment create value? For instance, through what mechanisms does it convert resources (e.g., funding, expertise, relationships) into significant value?
- What ways of working and other critical factors enable maximum value from the investment?

A value proposition focuses on the essence of value – its real or potential merit, worth or significance to people and groups – and how this value is produced. This goes beyond the process of change to consider what is valued about the intervention and how this value emerges (King, 2021).

Theories of change and value propositions are important because they provide critical points of reference for structuring causal and evaluative claims. They also assist in identifying context-specific criteria of VfI, step 2 of the VfI process.

Step 2 – Criteria

VfI criteria are the aspects of 'good resource use' evaluated in VfI assessments. Criteria should be defined contextually (Schwandt, 2015) and collaboratively with stakeholders (King et al., 2013) to ensure they are meaningful and specific to a policy, portfolio or programme – so a definitive set of criteria cannot be specified here. The "Five Es" used by organisations like FCDO (DFID, 2011) and NAO (2024) are a common example. The generic definitions below can be expressed in context-specific terms (King et al,. 2023).

- **Economy**: Good stewardship of resources to purchase inputs.
- **Efficiency**: Productive use of inputs to maximise outputs.
- **Effectiveness**: Achieving outcomes and impacts; meeting needs.
- **Cost-effectiveness**: Creating enough value to justify the resources invested.
- **Equity**: Addressing inequities and improving distributive justice through programme design, delivery, and outcomes.

These criteria are not exhaustive. Additional aspects such as ethics, relevance, coherence, and sustainability may also define good resource use (King, 2019; King, 2023i). When departments are running multiple programmes aimed at contributing to different aspects of an overarching problem, criteria can be developed to apply not just to one intervention but to a whole portfolio.

The Five Es are more suited to *ex-post* than *ex-ante* VfM assessment, as there are usually too many unknowns to assess economy, efficiency and effectiveness of a potential future investment. A leading example of a set of *ex-ante* criteria is the *Green Book's* five case model for business cases (strategic

case, economic case, commercial case, financial case, management case). These cases tend to be developed by different people with different sets of expertise, and would require coordination to engage with case leads and develop a coherent *ex-ante* VfM assessment. Applying explicit evaluative reasoning would enhance the application of the five-case model.

Step 3 – Standards

VfI standards are defined levels of performance that apply to each criterion. The standards specify what the evidence would look like at different levels of performance (Davidson, 2005). They articulate an agreed basis for making judgements. For example, King et al. (2023, p. 25) provide generic standards defining four levels of performance:

- **Excellent**: Meeting all reasonable expectations and/or targets bearing in mind context, and substantively exceeding some. Room for incremental improvements.
- **Good**: Generally meeting reasonable expectations/targets, allowing for a few minor exceptions. Some improvements may be needed.
- **Adequate**: Not meeting all reasonable expectations/targets, but fulfilling minimum 'bottom-line' requirements and showing acceptable progress overall. Significant improvements may be needed.
- **Poor**: Not fulfilling minimum requirements or not showing acceptable progress. Immediate and major improvements are needed.

Standards do not have to use the labels 'excellent', 'good', 'adequate' and 'poor', nor are there always four levels. Standards are developed to be fit-for-purpose in the context in which they are to be used. While the generic standards above can be used 'out of the box', King et al. (2023) recommend defining programme-specific standards aligned with the generic definitions.

A matrix of criteria and standards is called a rubric. King et al. (2023) provide guiding principles for developing rubrics with stakeholders including determining who to include, using participatory processes, suspending conversations about measurement, tailoring rubrics to the specific context of a policy, portfolio, or programme, and keeping them simple.

Step 4 – Evidence needed

The content of the rubrics helps in determining the types and sources of evidence needed to address the agreed criteria. This step also involves deciding on the design and mix of methods for collecting and analysing evidence, including appropriate approaches to causal inference. Method selection should be contextual and negotiated with stakeholders (Montrosse-Moorhead, Griffith, & Pokorny, 2014; Patton & Campbell-Patton, 2021; Yarbrough et al., 2011).

Determining the necessary evidence involves examining the criteria and standards and asking questions such as: What evidence is needed and will be credible to address the criteria? Credible to whom? What kinds of evidence are feasible in this context? What different sources of data are available for triangulation? In cases where evidence is lacking, what information could provide approximate answers to important questions? (McKegg et al., 2018).

In a Value for Investment (VfI) evaluation, the inclusion of economic evaluation methods (such as cost-benefit analysis, cost-effectiveness analysis, or cost-utility analysis) should be carefully considered (King, 2023c; 2023d). As noted previously, economic methods can contribute valuable insights but may be insufficient for a comprehensive evaluation of a complex social programme.

CBA should be considered for inclusion within a VfI evaluation when:

- The cost-benefit test (*Kaldor-Hicks efficiency* – i.e., is society better off overall?) is relevant;
- Outcomes and costs can meaningfully be quantified, attributed to the intervention and monetised;
- The necessary data, resources, time and skills are available to conduct the analysis; *and*
- The inclusion of CBA is ethical and responsive to stakeholders, in line with programme evaluation standards (King, 2019; 2023a; 2023b; UKES, n.d.; Yarbrough et al., 2011).

If these conditions are *not* met, then CBA may *not* be the right method for the circumstances.

CBA should be combined with other methods when:

- The cost-benefit test isn't the only relevant criterion (e.g., depending on context it may be important to balance additional considerations such as equity, accessibility, acceptability, sustainability, ethics, etc);
- There are material benefits or costs that can't be accurately and credibly monetised;
- Other forms of evidence (quantitative and/or qualitative) are needed to address evaluation questions and criteria;
- Tensions or disparities exist between groups, requiring disaggregation of values;
- Collective sense-making and deliberation are desirable for evaluation validity, credibility and utility;
- Reducing the value of a programme or policy to a single indicator risks oversimplifying complex issues;
- A consequentialist perspective is insufficient (e.g., where processes or procedural fairness matter) (King, 2019).

In our experience, one or more of these circumstances is likely to apply in the vast majority of social investments. Combining CBA with other evaluation methods allows the VfM assessment to take a wider view of resources, actions, impacts and value than CBA can do on its own.

In a mixed methods evaluation, there are opportunities to design an overall VfM assessment that ensures complementarity and coherence between methods. To illustrate: the design of a CBA could be informed by a literature review and stakeholder interviews; the inclusion of qualitative valuing approaches (such as rubrics) could influence the scope of a CBA by permitting a narrower-scope CBA that contributes a sound estimate of NPV based on monetisable factors without stretching the limits of credibility by attempting to include the hardest-to-measure intangible values (King, 2024c). VfI also offers a viable alternative to CBA in circumstances where CBA isn't feasible or appropriate (King, Crocket, & Field, 2023).

The evaluation design phase culminates in the preparation of a written evaluation framework and plan. Once this document is endorsed by end-users and stakeholders, the remaining steps of the evaluation can proceed.

3.2.2. Implementing the evaluation

Steps 5-8 of the VfI process involve: gathering and analysing evidence; bringing the streams of evidence together, guided by the rubrics, to synthesise and make evaluative judgements; and reporting findings.

Step 5 – Gather evidence

Gathering the necessary evidence involves adhering to accepted good practices and ethical standards associated with the selected methods. Some evidence may be collected directly by the evaluation team, while other evidence may come from existing sources (McKegg et al., 2018).

Depending on the situation, the evidence could include a mix of quantitative and qualitative data, gathered through methods such as observation, measurement, surveys, focus groups, interviews, administrative databases, financial accounts, literature, and documentation. Where feasible and appropriate, the methods can include economic analysis.

The criteria and standards provide a structured framework for data collection, enabling efficient and focused systems to organise the evidence for analysis. For instance, the design of surveys and interview guides can systematically follow the content of the rubrics, using sub-headings consistent with the sub-criteria, ensuring that the feedback directly addresses the criteria (King, Crocket, & Field, 2023).

Step 6 – Analysis

During this step, each stream of evidence is analysed individually, to identify findings that are relevant to the evaluation questions, criteria, and standards. It is in this step that a CBA or other economic evaluation, if included, would be undertaken, following *Green Book* and *Magenta Book* guidance.

This step also involves addressing the question of causality or contribution. In other words, what is the impact of the policy or programme *over and above* what would have been achieved under a status quo scenario where no policy reform was implemented. While CBA requires a counterfactual scenario including quantitative measures of benefits and costs, VfI can accommodate any approaches to causal inference or contribution analysis, including theory-based and additionality approaches (King et al., 2023).

Step 7 – Synthesis and judgements

Synthesis is a distinct and separate step from analysis. While analysis involves examining pieces of evidence individually, synthesis combines all the evidence to determine what it shows and what it means overall. This is also the stage where evaluative judgements are made, interpreting the evidence through the lens of the criteria and standards (McKegg et al., 2018).

Stakeholder engagement is crucial during these steps, because a collaborative sense-making process enhances evaluation validity by incorporating multiple perspectives. Additionally, it fosters stakeholders' understanding, ownership, and use of the evaluation findings (King, 2023o).

A rubric can be used to support judgements about findings from CBA alongside other findings. An example is provided below (King, 2024a). Section 7.2 provides more information about break-even analysis.

Figure 2: Generic standards and application to break-even analysis

	Generic definitions (King, Wate, Namukasa, Hurrell, Hansford, Ward, & Faramarzifar, 2023)	**Application to break-even analysis** (King, 2024)
Excellent	Meeting or exceeding all reasonable expectations/targets bearing in mind context. Room for incremental improvements.	Breaks even beyond a reasonable doubt.
Good	Generally meeting reasonable expectations/ targets, allowing for minor exceptions. Some improvements needed.	Breaks even on balance of probabilities.
Adequate	Not meeting expectations but fulfilling minimum requirements and showing acceptable progress overall. Significant improvements needed.	Credible prospect of breaking even.
Poor	Not fulfilling minimum requirements; unacceptable progress. Urgent improvements needed.	Little or no prospect of breaking even.

Step 8 – Reporting

The objective of reporting is to communicate findings, presenting a clear and accurate account of evaluative judgements, supported by evidence and transparent reasoning (King, 2019). A good evaluation report gets straight to the point and gives a clear statement of findings, up front, including VfI ratings (excellent, good, etc) against the criteria (King et al., 2023). Examples are provided in the next section.

4. Examples

The following examples illustrate the use of the VfI approach.

4.1. VfI evaluation that included CBA within a mix of methods

This series of social and economic impact assessments for the International Atomic Energy Agency (IAEA) illustrate the use of the Value for Investment approach including mixed methods and rubrics designed collaboratively with stakeholders. CBA, including break-even analysis, was combined with surveys, case studies, key expert interviews, and administrative data – providing insights that would not have been gained from the use of CBA alone. The reports were structured around key evaluation criteria, with concise findings up front supported by evidence and transparent reasoning, and full details in a series of annexes.

A worked example is provided here:

King, J. (2022). *A value for investment report deconstructed*. Web page.

The reports are:

King, J., McKegg, K., Arau, A., Schiff, A., Garcia Aisa, M. (2020). *Social and Economic Impact Assessment of Mutation Breeding in Crops of the RCA Programme in Asia and the Pacific.* International Atomic Energy Agency, Vienna.

King, J., Arau, A., Schiff, A., Garcia Aisa, M., McKegg, K. (2022). *Social and Economic Impact Assessment of the RCA Programme: Radiotherapy Case Study*. International Atomic Energy Agency, Vienna.

King, J., Arau, A., Schiff, A., Garcia Aisa, M., McKegg, K. (2022). *Social and Economic Impact Assessment of the RCA Programme: Non-Destructive Testing Case Study*. International Atomic Energy Agency, Vienna.

4.2. VfI evaluation that provided a viable alternative to CBA

This example from New Zealand shows how the Value for Investment approach can provide answers to VfM questions in circumstances where CBA isn't feasible. This evaluation of the national rollout of a Youth Primary Care Mental Health and Addictions initiative was conducted early in the life of the services and before quantitative outcome measures were implemented. The mix of methods used in this evaluation included surveys, interviews, documents review, and analysis of service output data.

The evaluation report is available here:

Field, A., Crocket, A., Garden, E., King, J., Moss, M., Parslow, G., Schiff, A., Spee, K., Wehipeihana, N. (2023). *Youth Primary Mental Health and Addictions Evaluation*. Final Report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Limited, Auckland.

VfI application and insights report: In addition, the evaluation team was commissioned to prepare a VfI guide, providing a detailed account of the application of VfI's eight steps and the participatory processes involved, providing a road map for future evaluations and capacity building in the sector:

King, J., Crocket, A., Field, A. (2023). *Value for Investment: Application and Insights*. Youth Primary Mental Health and Addictions Evaluation. Exemplar report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Ltd.

5. Conclusion

In this Guide, we introduce a significant advancement in the assessment of VfM by integrating the complementary strengths of economics and evaluation.

CBA contributes to VfM assessment, providing insights that would otherwise be difficult to gain. In particular, systematic and rational analysis of costs and benefits, valued in commensurable units, enables the value of a policy or programme to be reduced to a single indicator, communicating value in a straightforward and compelling way.

However, when it comes to evaluating VfM of social investments, CBA is usually insufficient to fully answer VfM questions. For example, VfM assessment often involves balancing multiple criteria, whereas CBA assesses performance against a single criterion. The scope of a CBA may be constrained by what is measurable and monetisable, while in social investments some of the most valuable outcomes may be the hardest to value monetarily. Value for Investment offers a way forward.

5.1. Advantages of Value for Investment

By integrating CBA (and other economic methods and concepts) within a broader evaluative framework that incorporates multiple values and diverse evidence sources, the inter-disciplinary VfI approach offers a more comprehensive assessment of policies and programmes, drawing on the strengths of economic evaluation without limiting the evaluation to economic metrics. The participatory and mixed-methods nature of this approach ensures it is both practical and inclusive, facilitating stakeholder engagement and robust evaluative reasoning. Value for Investment evaluations can:

- Accommodate multiple and diverse criteria;
- Incorporate value in multiple ways (monetary, cardinal, ordinal, and qualitative);
- Incorporate 'process value' (e.g., evaluating efficiency and other factors from the middle section of the theory of change) in addition to costs and consequences;
- Accommodate mixed methods (qualitative and quantitative evidence) and facilitate contextual selection of methods without prescribing or privileging certain methods over others;
- Promote transparency in the way evaluative judgements are made and reported;
- Facilitate comparable VfM assessments by developing and using consistent criteria, standards and metrics for multiple investments across a portfolio;
- Limit the contribution of CBA to providing a sound estimate of attributable and monetisable benefits and costs, leaving intangible factors to the wider mix of methods – a position that accords with the arguments of Sunstein (2018) and Adler & Posner (2006);
- Support VfM assessments that are inclusive, responsive, contextually viable and meaningful for stakeholders whose lives are affected by the policies and programmes (King, 2019; Yarbrough et al., 2011).

As demonstrated through its application in real-world contexts, the VfI approach not only enhances the rigour of VfM assessments but promotes informed and evidence-based policymaking.

5.2. Challenges in using Value for Investment

As with any approach, there are challenges in the use of VfI. A key issue is ensuring consistency and comparability across different VfM assessments. When criteria are tailored to each specific context, it becomes more difficult to compare or rank different programmes, whereas CBA can (at least in principle) support like-with-like comparisons across diverse alternatives. Nonetheless, portfolio-specific rubrics can be developed that enable comparisons across a group of investments, with standardised economic and other metrics to promote comparability (King, 2019).

Moreover, to promote consistency to the extent possible, each set of bespoke rubrics can be aligned conceptually with generic definitions of criteria (such as the five Es) and standards (such as those shown in Figure 2) so that the underlying definition of each rubric is broadly consistent while context-specific details may vary (King et al., 2023).

A comprehensive VfM assessment using the VfI approach can be a resource-intensive undertaking. However, VfI evaluations need not be time consuming or expensive. The four guiding principles (inter-disciplinary, mixed methods, evaluative reasoning, and participatory) can be implemented to different levels of comprehensiveness as appropriate to provide proportionate assessments of appropriate quality in different circumstances. For example, criteria and evidence can be prioritised (King, 2023m).

Stakeholder bias toward efficiency indicators, like NPV, can hinder the adoption of VfI. The simplicity of summarising programme outcomes and costs in a single indicator is appealing. However, NPV alone cannot provide a comprehensive VfM assessment. In contrast, making evaluative judgements based on multiple criteria and diverse evidence requires a more nuanced and detailed presentation, although these findings can still be summarised in a clear statement, such as 'this programme meets the threshold for good VfM as defined by the agreed criteria and standards' (King, 2019).

CBA may appear more objective than VfI to some stakeholders – but this is an issue of buy-in, not validity. Both VfI and CBA, like all forms of evaluation, require judgement. The conclusions reached in a CBA are affected by analyst judgements about: what scope of benefits and costs are relevant, material and feasible to include; the perspective of the study; the time horizon over which benefits and costs are likely to last or matter; the discount rate; approaches to monetary valuation; and the interpretation and presentation of results. In a VfI evaluation, judgements are transparent and reached with stakeholders (King, 2023o). Evaluative reasoning, based on agreed criteria and standards, provides a transparent and systematic approach to gathering, analysing, and synthesising evidence. Explicit evaluative reasoning guards against personal subjectivity by articulating a shared set of values and making the evaluation process traceable and challengeable (King, 2019, p. 209). Ultimately, both CBA and VfI provide valid perspectives, making appropriate use of robust evidence together with human judgement.

5.3. Bottom-line

The principles and process of the Value for Investment approach are based on sound theory and practice from evaluation and economics. Value for Investment isn't another method – it's a system to guide the selection and contextually-appropriate use of existing evaluation methods to answer value-for-money questions. For this reason, it is applicable to any setting where the objective is to address evaluative questions about resource use and value creation.

Whether you're doing a CBA or a VfI evaluation, we recommend involving evaluators early – including during programme design – so that the right data and information collection processes are established from the outset.

6. For more information

6.1. Value for Investment training

VfI training workshops are jointly offered to staff and consultants of the UK public sector by Verian and Julian King & Associates. Training workshops can be provided online or in-person and can be customised to meet needs, typically ranging from three hours to two days duration. Workshops can be scheduled on request.

VfI training workshops for evaluators are also offered periodically through the UK Evaluation Society (UKES), through a collaboration between Oxford Policy Management, Julian King & Associates, and Verian.

The VfI approach is included in evaluation training at whole-of-government level in the UK through the Evaluation Task Force (ETF), a joint HM Treasury and Cabinet Office unit established to ensure evidence and evaluation sit at the heart of spending decisions. The ETF's value for money training module recommends the use of the VfI approach as set out in King & OPM (2018), and now supplemented in key areas by this Guide.

6.2. Value for Investment resources

The following documents and websites are recommended for more information and complement this Guide. A full reference list is provided at the end of this document.

King, J., Wate, D., Namukasa, E., Hurrell, A., Hansford, F., Ward, P., Faramarzifar, S. (2023). *Assessing Value for Money: the Oxford Policy Management Approach*. Second Edition. Oxford Policy Management Ltd.

King, J., Crocket, A., Field, A. (2023). *Value for Investment: Application and Insights*. Youth Primary Mental Health and Addictions Evaluation. Exemplar report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Ltd.

Julian King & Associates Limited: www.julianking.co.nz

Evaluation and Value for Investment on Substack: https://juliankingnz.substack.com

6.3. Contact us

Alex Hurrell UK Head of Evaluation, Verian London email alex.hurrell@veriangroup.com	Julian King Director, Julian King & Associates Limited Auckland, New Zealand email jk@julianking.co.nz

7. Annexes

7.1. Sensitivity analysis and scenario analysis

Typically, conducting a CBA is a modelling exercise. Usually, we don't have precise measurements of all the values we need (the exception would be an *ex-post* assessment of a financial investment or a programme with clearly attributable and monetisable impacts). Most of the time, CBA involves modelling costs and benefits, informed by available programme data together with external evidence and transparent assumptions (King, 2023k).

Estimates and assumptions may be needed at every step – for example, determining: which costs and benefits are material, attributable and monetisable; their quantities; their monetary valuations; their timing; how far into the future they will last or matter; what discount rate to use and other factors. These various assumptions and pieces of data are held together by a mathematical logic that involves more assumptions and estimates about the relationships between causal and consequential variables (King, 2023k).

This central role of estimates and assumptions in CBA has important implications when interpreting the results of a CBA. Usually, the output of a CBA isn't a prediction or a definitive answer – it is an estimated value of a scenario that we can't directly measure or predict but deem plausible. One of the most useful functions CBA can perform is to illuminate the implications of uncertainty (through sensitivity analysis) and produce a range of estimates for a range of different plausible futures (through scenario analysis) (King, 2023k).

Sensitivity analysis tests individual assumptions in isolation, to determine the extent to which model outputs (like NPV) vary in response to changes in the assumed value of an input variable. To illustrate, imagine we are conducting an *ex-ante* evaluation of a proposed intervention to help people quit smoking. We are reasonably confident in our estimate of what the programme is likely to cost. Outcomes are uncertain but can be estimated from evaluations of similar interventions in other countries, together with local data such as the burden of disease and premature death attributed to smoking in our country.

In the first instance, we can set our modelling assumptions to 'most likely' ballpark values and produce a *point estimate* (a single value), e.g., a NPV of £100 million. Using CBA to produce point estimates is valid. However, to provide transparent analysis of the uncertainty behind the point estimate, it is a good practice to conduct sensitivity analysis.

In this illustrative case, we need to make assumptions is the level of effectiveness – what proportion of participants will successfully quit smoking. Findings from overseas evaluations may give some indication of a realistic level of effectiveness, but we can't predict exactly how effective this new programme will be. Subjecting the level of effectiveness to sensitivity analysis involves running the model for a range of lower and higher effectiveness levels to understand now these variations affect the results of the analysis. For example, we might find that when we reduce the assumed level of effectiveness by ten percentage points that the NPV reduces by 20 percent, to £80 million – i.e., we can say that the NPV is sensitive to the assumed level of effectiveness.

Sometimes, a small change in an assumption can have a large impact on the NPV. In other cases, a large change in an assumption has a small impact on the NPV. Sensitivity analysis can highlight which assumptions are the weakest links in our model (King, 2023k).

Scenario analysis takes a 'what-if' approach, changing combinations of assumptions for a range of optimistic and pessimistic scenarios, to identify a range of plausible results. For example, perhaps our

sensitivity analyses reveal that the NPV is sensitive to assumptions about the level of effectiveness, the participation rate, and the cost per participant. By adjusting all three variables together, we find the NPV varies from £60 million under a pessimistic set of assumptions to £140 million under a more optimistic scenario. These *range estimates* (£60-140 million) are, in our view, a more honest and informative way of reporting potential future values than only giving a point estimate.

A more sophisticated approach to scenario analysis, *Monte Carlo simulation* involves estimating probability distributions for uncertain variables and creating a large number of scenarios (say, 1000), changing the input variables each time according to their probability distributions together with a randomisation process. The resultant output shows the probability distribution of the NPV, which can be plotted on a frequency graph.

Running a Monte Carlo simulation can help in judging how risky a positive outcome may be. For example, a simulation might indicate that 760 out of 1,000 scenarios have a NPV greater than zero. Whether that's an encouraging or discouraging result would depend on the risk appetite of decision-makers, but the insight helps us make an informed decision (King, 2023k).

7.2. Break-even analysis with a rubric

A full cost-benefit analysis isn't always feasible. In summative evaluations of policies and programmes, it may be desirable to include CBA within our mix of methods but real-world challenges such as missing data, time or budget constraints can sometimes preclude a comprehensive CBA. For example, we may have clear information about costs but face uncertainties related to identifying, quantifying, attributing and/or monetising benefits. These challenges can be addressed but doing so requires specialist skills, and a comprehensive CBA in these situations can be time-consuming and expensive.

Break-even analysis offers a way through. Break-even analysis enables a rapid judgement, using minimal data, of the prospect of benefits equalling costs, without the need for a full CBA (King, 2024a). A typical question that we would set out to answer through CBA is along the lines of: *How does the present value of benefits compare to the present value of costs?* To answer this question, we would need sound estimates of both benefits and costs. Break-even analysis eases the requirements a little, by instead asking: *What would we have to assume for benefits to equal costs?* or, *Can we find enough benefits to equal costs?* It is often feasible to answer these questions when there isn't enough data for a comprehensive CBA.

In Figure 2, we provide a rubric that can be used to support the process of making evaluative judgements in break-even analysis (King, 2024a). This rubric borrows two legal concepts to differentiate levels of performance: 'beyond a reasonable doubt', and 'on balance of probabilities'. Examples of what the evidence might look like at the four different levels of the rubric include:

- **Excellent (breaks even beyond a reasonable doubt):** Results are well ahead of the level of effectiveness required to break even; modelling indicates that benefits exceed costs under any combination of plausible assumptions including pessimistic scenarios.

- **Good (breaks even on balance of probabilities):** The investment is more likely than not to break even; results are a little ahead of the level of effectiveness required to break even; modelling indicates that break-even is likely under realistic mid-range scenarios.

- **Adequate (a credible prospect of breaking even):** Results are in the right ballpark to break even; modelling indicates break-even could be achieved under at least some plausible combinations of assumptions.

- **Poor (little or no prospect of breaking even):** Results are well below the level of effectiveness required to break even; modelling indicates unreasonably optimistic assumptions are required to break even (King, 2024a).

The rubric definitions above should be regarded as illustrative examples. Rubrics should be developed with stakeholders to ensure they fit the context. The hurdles may be set higher or lower depending on stakeholders' expectations. For example:

- For some investments, a 'credible prospect of breaking even' might not be good enough to qualify as Adequate. Risk appetite is contextual.

- The time horizon for breaking even should reflect how long the benefits and costs are expected to last or matter. For example, for a relapsing-remitting condition like addiction, remaining abstinent for a few years might be considered a good result, whereas an infrastructure investment such as a new hospital would be expected to last for multiple decades.

- The discount rate is contextual. While the *Green Book* specifies a discount rate of 3.5%, scenarios with higher and lower rates can be included. Typically, a riskier investment would justify a higher discount rate, while an intergenerational impact might justify a low or perhaps even a zero discount rate.

- Merely breaking even might not be enough. To illustrate, in a particular context we may set a minimum expectation (Adequate) of benefits exceeding costs by a ratio of (say) 2:1 and we may expect the best investments (Excellent) to have a benefit-to-cost ratio (BCR) of 10:1 or more.

- We can add a 'plus-or-minus' factor to take account of intangible benefits or costs. For example, if our judgement is on the line between, say, Adequate and Poor, and we know that there are significant intangible benefits not included in the NPV, we could take that into account as rationale for erring on the side of an Adequate rating (King, 2024a).

7.3. What about social return on investment (SROI)?

CBA and SROI share an objective of assessing a policy, programme or project's value relative to its costs. They both involve identifying, quantifying, attributing, monetising, discounting, and aggregating costs and benefits. Both CBA and SROI can be applied at any scale and from either macro (whole-of society) or micro (programme or project) perspectives. Either method can be used *ex-ante* (to inform decisions) or *ex-post* (to evaluate outcomes). In practice, CBA is widely used for *ex-ante* assessments. SROI is popular in philanthropy and social impact sectors, while CBA is favoured by for-profit and government decision-makers (King, 2023p).

Both methods produce similar indicators (such as NPV and BCR) but they approach the task from different perspectives and emphasise different aspects of good practice. CBA focuses on methodological precision, accuracy, and reliability, addressing attribution, valuation, discounting, and dealing with uncertainty. In relative terms, SROI places greater emphasis on stakeholder engagement and supplementing quantitative analysis with qualitative description of value (King, 2024b).

SROI, emerging in the mid-1990s from social accounting, contrasts with CBA's roots in welfare economics, tracing back to the 18th and 19th centuries. CBA has a well-established professional community, extensive training programmes, and a comprehensive body of literature. SROI has helped to make systematic analysis of benefits and costs more accessible to non-economists which is both a positive (supporting more widespread use) and carries some risks (due to the variable training

and skill levels of practitioners). A recent article (Siegal, 2022) highlighted some common deficiencies seen in SROI studies, including variable reliability in attributing benefits and valuing impacts, and vulnerability to false precision in NPV or BCR estimates. These issues reflect practice variability rather than inherent flaws in SROI.

We contend that combining social CBA's methodological rigour and SROI's attention to stakeholders could enhance the application of both methods (King, 2024b). Ultimately, any systematic analysis of benefits and costs should be applied with attention to context, maintaining methodological rigor, and stakeholder involvement. Moreover, a comprehensive evaluation may require integrating either CBA or SROI within a VfI framework.

8. References

Adler, M.D., Posner, E.A. (2006). *New Foundations of Cost-Benefit Analysis.* Harvard University Press. ISBN 9780674022799.

Davidson, E.J. (2005). *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation.* Sage.

DFID (2011). *DFID's Approach to Value for Money (VfM).* UK Department for International Development.

Field, A., Crocket, A., Garden, E., King, J., Moss, M., Parslow, G., Schiff, A., Spee, K., Wehipeihana, N. (2023). *Youth Primary Mental Health and Addictions Evaluation.* Final Report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Limited, Auckland.

Flyvbjerg, B., Bester, D.W. (2021). The Cost-Benefit Fallacy: Why cost-benefit analysis is broken and how to fix it. *Journal of Benefit-Cost Analysis,* Oct, 1025, doi 10.1017/bca.2021.9.

Fournier, D.M. (1995). Establishing Evaluative Conclusions: A distinction between general and working logic. In D.M. Fournier (Ed). *Reasoning in Evaluation: Inferential Links and Leaps. New Directions for Evaluation* 58, 15-32.

Funnell, S.C., Rogers, P.J. (2011). *Purposeful Program Theory: Effective use of theories of change and logic models.* Wiley.

Gargani, J., King, J. (2023). Principles and methods to advance value for money. *Evaluation,* 1-19. DOI: 10.1177/13563890231221526.

Greene, J.C. (2007). The generative potential of mixed methods inquiry. *International Journal of Research & Method in Education,* 28(2).

HM Treasury. (2022a). *The Green Book: Central government guidance on appraisal and evaluation*. Government of the United Kingdom. Open access guidance and resources.

HM Treasury. (2022b). *The Magenta Book: HM Treasury guidance on what to consider when designing an evaluation*. Government of the United Kingdom. Open access guidance and resources.

Julnes, G. (2012). Promoting valuation in the public interest. In G. Julnes (Ed). *Promoting Valuation in the Public Interest: Informing Policies for Judging Value in Evaluation*. *New Directions for Evaluation,* 133, pp 109-129

Kahneman, D. (2012). *Thinking, Fast and Slow.* Penguin Press. ISBN 9780141033570

King, J. (2017). Using Economic Methods Evaluatively. *American Journal of Evaluation*, *38*(1), 101–113.

King, J. (2019). *Evaluation and Value for Money: Development of an approach using explicit evaluative reasoning.* (Doctoral dissertation). Melbourne, Australia: University of Melbourne.

King, J. (2021). Expanding theory-based evaluation: incorporating value creation in a theory of change. Evaluation and Program Planning, Volume 89, December 2021.

King, J. (2022). *A value for investment report deconstructed.* Web page.

King, J. (2023a). How should Program Evaluation Standards inform the use of cost-benefit analysis in evaluation? *Journal of MultiDisciplinary Evaluation.* Vol. 19, No. 43.

King, J., Allan, S. (2018). Applying Evaluative Thinking to Value for Money: The Pakistan Sub-National Governance Programme. *Evaluation Matters—He Take Tō Te Aromatawai*, 4, pp. 207-235.

King, J., McKegg, K., Arau, A., Schiff, A., Garcia Aisa, M. (2020). *Social and Economic Impact Assessment of Mutation Breeding in Crops of the RCA Programme in Asia and the Pacific.* International Atomic Energy Agency, Vienna.

King, J., Arau, A., Schiff, A., Garcia Aisa, M., McKegg, K. (2022). *Social and Economic Impact Assessment of the RCA Programme: Radiotherapy Case Study*. International Atomic Energy Agency, Vienna.

King, J., Arau, A., Schiff, A., Garcia Aisa, M., McKegg, K. (2022). *Social and Economic Impact Assessment of the RCA Programme: Non-Destructive Testing Case Study*. International Atomic Energy Agency, Vienna.

King, J., Crocket, A., Field, A. (2023). *Value for Investment: Application and Insights.* Youth Primary Mental Health and Addictions Evaluation. Exemplar report for Te Whatu Ora – Health New Zealand. Dovetail Consulting Ltd.

King, J., Guimaraes, L. (2016). Evaluating Value for Money in International Development: The Ligada Female Economic Empowerment Program. *eVALUation Matters*, Third Quarter, 2016, pp. 58-69. Africa Development Bank.

King, J., McKegg, K., Oakden, J., Wehipeihana, N. (2013). Rubrics: A method for surfacing values and improving the credibility of evaluation. *Journal of MultiDisciplinary Evaluation,* 9(21), 11-20.

King, J., Namukasa, E., Hurrell, A. (2023). Incorporating value creation in a theory of change. *The Evaluator,* Spring, 6-9.

King, J. & OPM (2018). *OPM's approach to assessing value for money – a guide*. Oxford Policy Management Ltd.

King, J., Wate, D., Namukasa, E., Hurrell, A., Hansford, F., Ward, P., Faramarzifar, S. (2023). *Assessing Value for Money: the Oxford Policy Management Approach*. Second Edition. Oxford Policy Management Ltd.

Kinnect group & Foundation North. (2016). *Kia E Te Whakangao: Māori & Pacific Education Initiative Value for investment evaluation report*. Auckland: Foundation North.

McKegg, K., Oakden, J. Wehipeihana, N., King, J. (2018). *Evaluation Building Blocks: A Guide*. The Kinnect Group.

Montrosse-Moorhead, B., Griffith, J.C., & Pokorny, P. (2014). House with a view: Validity and evaluative argument. In J.C. Griffith & B. Montrosse-Moorhead (Eds). *Revisiting truth, beauty, and justice: Evaluating with validity in the 21st Century. New Directions for Evaluation,* 142, 95-105.

NAO (2024). *Assessing Value for Money.* National Audit Office. Available online.

Oakden, J., King, J. (2018). Evaluation, in M. Tolich & C. Davidson (Eds.). *Social science research in New Zealand: An introduction* (3rd ed; pp. 180-193). Auckland, New Zealand: Auckland University Press.

Patton, M.Q. & Campbell-Patton, C.E. (2021). *Utilization-Focused Evaluation* (5th Ed). Sage.

Schiff, A. (2024). *Data and information for economic evaluation of social investments.* Schiff Consulting. https://www.schiff.co.nz/en/resources/data-for-economic-evaluation/.

Scriven, M. (2013). *Key Evaluation Checklist.*

Siegal, K.M. (2022). The Tentative Promise of Social Return on Investment. *American Journal of Evaluation* 43(3).

Sunstein, C. R. (2018). *The Cost-Benefit Revolution*. Cambridge, MA: MIT Press.

UK Evaluation Society (n.d.). *Guidelines for Good Practice in Evaluation.*

Yarbrough, D.B., Shulha, L.M., Hopson, R.K., Caruthers, F.A. (2011). *The Program Evaluation Standards; A Guide for Evaluators and Evaluation Users (3rd Ed)*. Sage.

Wehipeihana, N. (2019). Increasing Cultural Competence in Support of Indigenous-Led Evaluation: A Necessary Step toward Indigenous-Led Evaluation. *Canadian Journal of Program evaluation, 34*(2), 368;384.

Evaluation and Value for Investment on Substack.

This Guide references a series of articles by Julian King, posted at: https://juliankingnz.substack.com. Full citations and links to specific articles are provided below.

King, J. (2023b). Contextually-responsive cost-benefit analysis. *Evaluation and Value for Investment.*

King, J. (2023c). Cost-effectiveness analysis: an evaluator's perspective. *Evaluation and Value for Investment.*

King, J. (2023d). Cost-utility analysis: an evaluator's view. *Evaluation and Value for Investment.*

King, J. 2023e). Cost-benefit analysis and the logic of evaluation. *Evaluation and Value for Investment.*

King, J. (2023f). Cost-benefit analysis through an evaluative lens. *Evaluation and Value for Investment.*

King, J. (2023g). Developing Rubrics with stakeholders. *Evaluation and Value for Investment.*

King, J. (2023h). Equity through a cost-benefit lens. *Evaluation and Value for Investment.*

King, J. (2023i). Evaluation vs. Value for Money: How VfM assessment differs (or not) from evaluation. *Evaluation and Value for Investment.*

King, J. (2023j). Evaluative thinking. *Evaluation and Value for Investment.*

King, J. (2023k). Modelling costs and benefits in uncertainty. *Evaluation and Value for Investment.*

King, J. (2023l). Objectivity and subjectivity in evaluation. *Evaluation and Value for Investment.*

King, J. (2023m). Rapid value for money assessment. *Evaluation and Value for Investment.*

King, J. (2023n). Rubrics and alternatives. *Evaluation and Value for Investment.*

King, J. (2023o). Sense-making with stakeholders and rubrics. *Evaluation and Value for Investment.*

King, J. (2023p). Social Return on Investment vs. Cost-Benefit Analysis. *Evaluation and Value for Investment.*

King, J. (2024a). Break-even analysis and rubrics. *Evaluation and Value for Investment.*

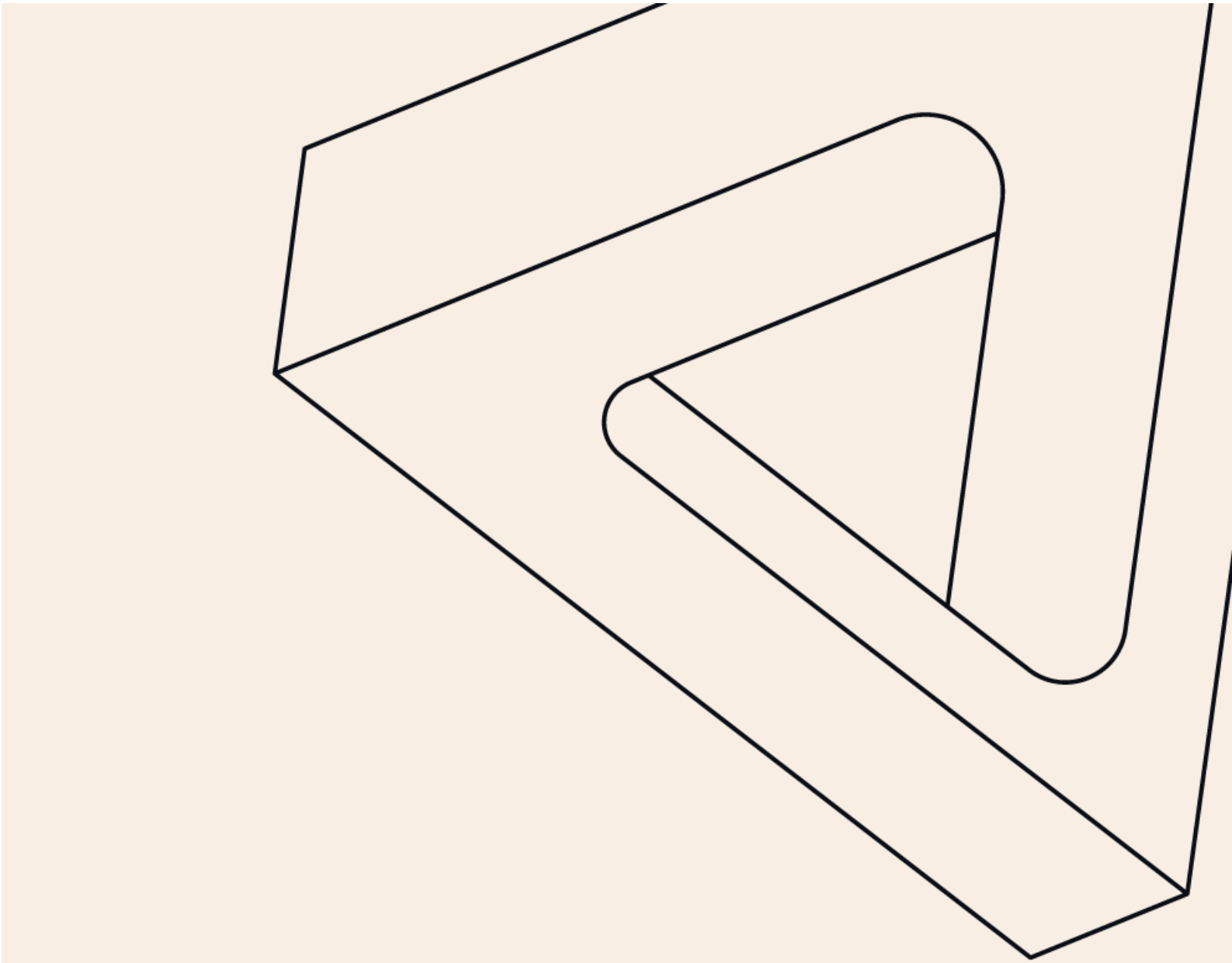
King, J. (2024b). CBA and SROI in the Garden of Eval. *Evaluation and Value for Investment.*

King, J. (2024c). Navigating CBA's triple tradeoff. *Evaluation and Value for Investment.*

King, J. (2024d). Wellbeing and cost-benefit analysis. *Evaluation and Value for Investment.*

verian

Powering decisions
that shape the world.



2024 Verian Group. All information and data contained in this document is the property of Verian Group and Julian King & Associates, and cannot be reproduced without our prior permission.

veriangroup.com