Evaluation and Value for Money

POPLHLTH 705 Evaluation Research Methods

University of Auckland

March 2020

Julian King

Introduction

Kia ora koutou katoa, I'm Julian King, an independent public policy consultant from Auckland. I'm a member of the Kinnect Group, a graduate and Honorary Fellow of the Centre for Program Evaluation, University of Melbourne, and an Associate of Oxford Policy Management. In this presentation I share an overview of an approach to evaluation and value for money (VFM) that I developed in my doctoral research. The approach is used globally to evaluate complex programs and policy reforms. I will also explain why we should use economic methods of evaluation more, and why we should combine them with other methods.

JK&A

Kinnect group

THE UNIVERSITY OF MELBOURNE

Oxford Policy Management

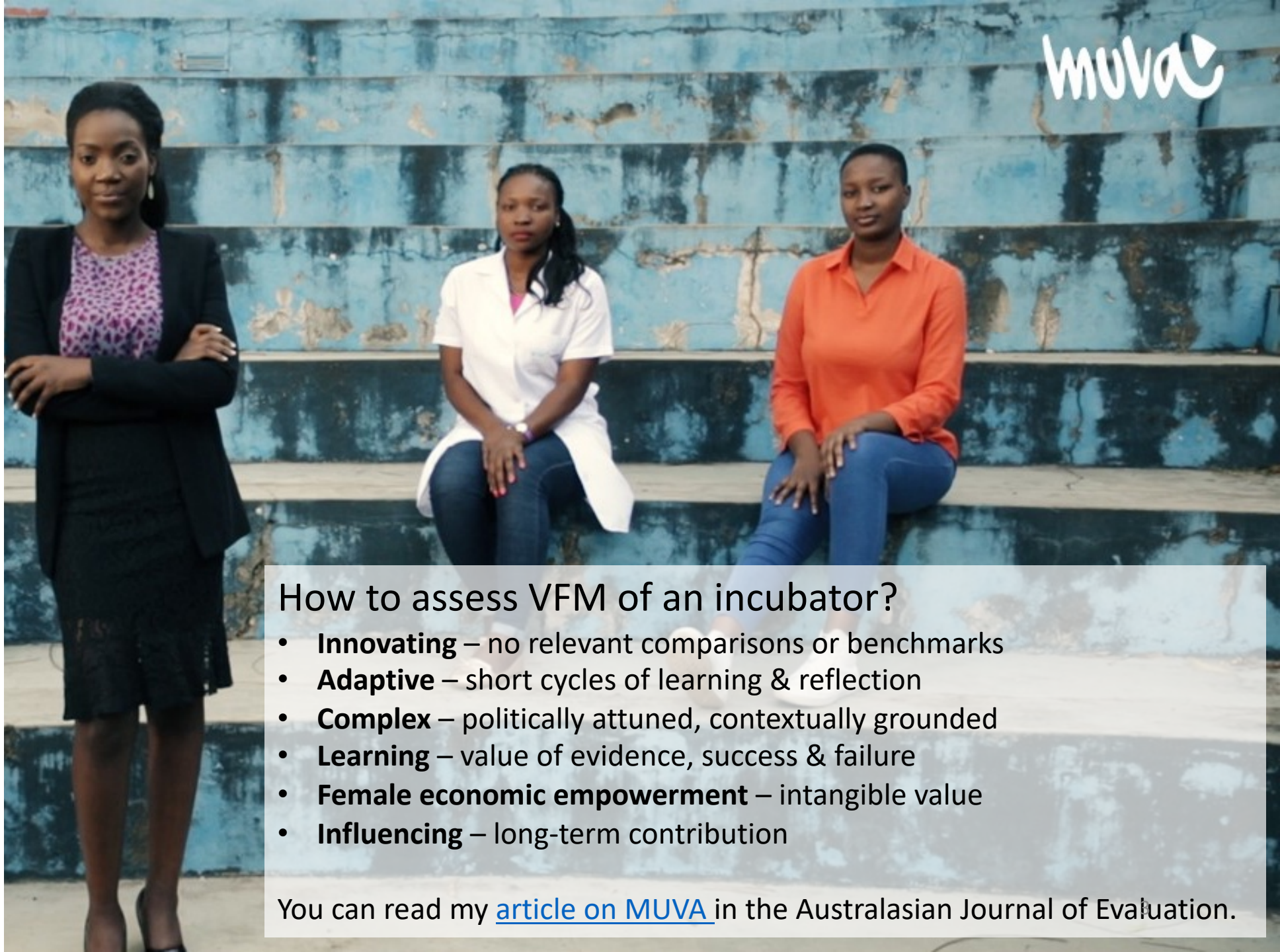
www.julianking.co.nz policy | evaluation | value for money

MUVA is a female economic empowerment program in urban Mozambique. It is an incubator, funded by the UK Department for International Development (DFID) and implemented by Oxford Policy Management (OPM). It develops and tests new approaches to female economic empowerment, generates evidence and learning, and influences other organisations to adopt and scale successful approaches. MUVA exemplifies many of the problems we encounter when assessing VFM in complex, adaptive programs.

UKaid
from the British people

Oxford Policy Management

www.muvamoz.co.mz

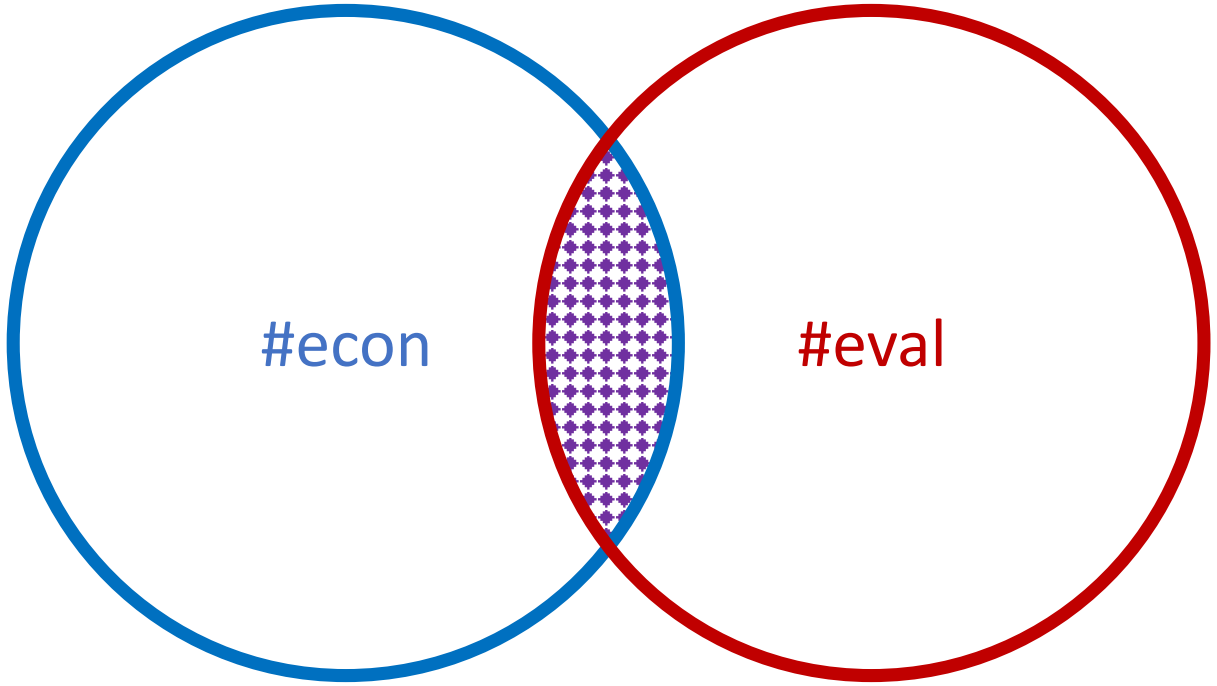


How to assess VFM of an incubator?

- **Innovating** – no relevant comparisons or benchmarks
- **Adaptive** – short cycles of learning & reflection
- **Complex** – politically attuned, contextually grounded
- **Learning** – value of evidence, success & failure
- **Female economic empowerment** – intangible value
- **Influencing** – long-term contribution

You can read my article on MUVA in the Australasian Journal of Evaluation.

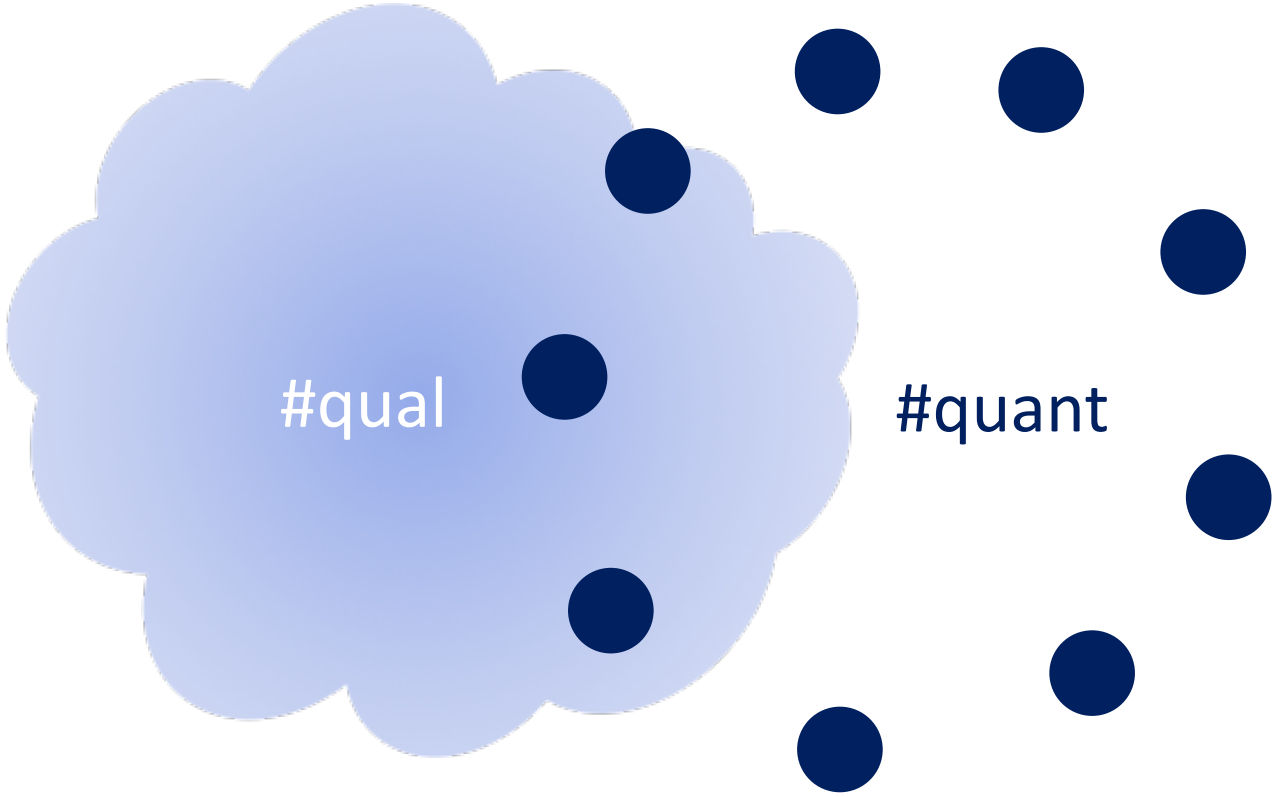
Inter-disciplinary



To help address these challenges, our approach to VFM assessment is inter-disciplinary. Evaluation and economics are both under-utilized in VFM assessment. Both disciplines offer useful frameworks and valuable insights – but neither discipline has all the answers. In our approach, we combine evaluative and economic thinking.

© www.julianking.co.nz

Mixed methods

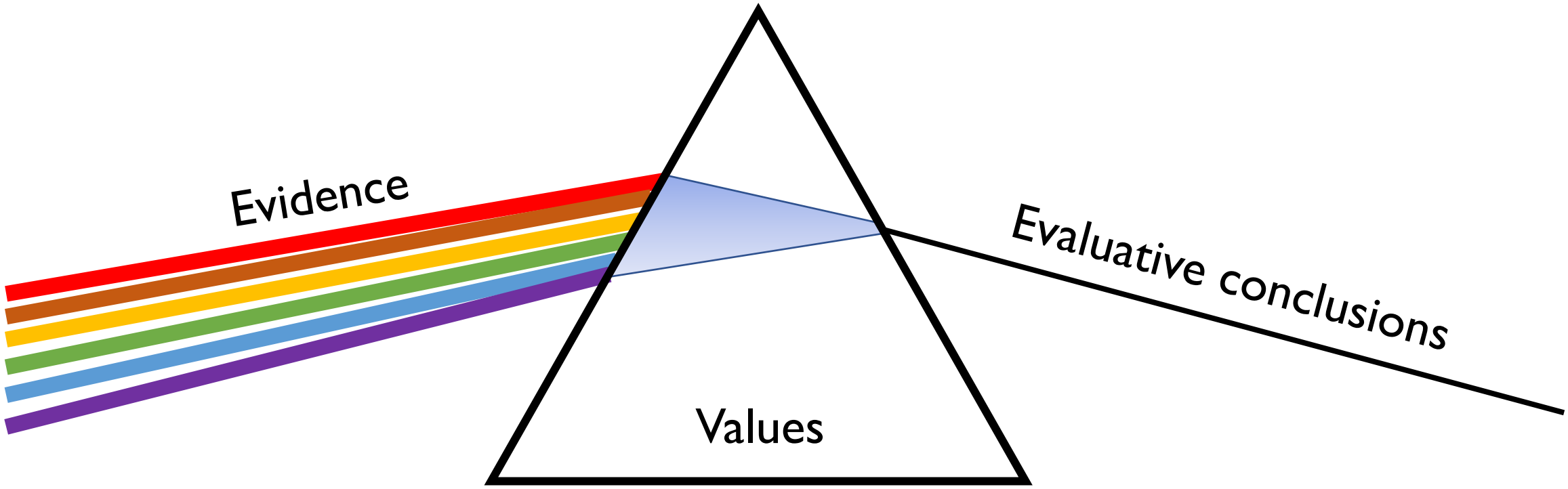


Our approach to VFM assessment uses mixed methods. We are always seeking to understand the story behind the numbers, and the best way I know to do that is to triangulate evidence from multiple sources – and to make sense of the evidence with stakeholders, rights-holders and end-users.

© www.julianking.co.nz

Evaluative reasoning

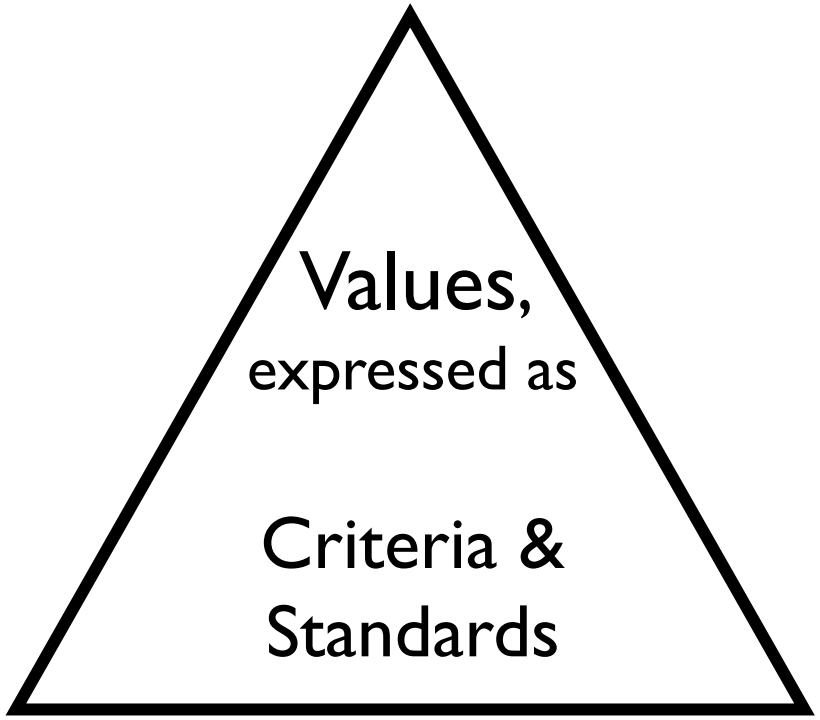
Evaluative reasoning provides the means for making transparent judgements from the evidence. A bit like a prism in reverse, values provide a lens for looking at a broad spectrum of evidence, and reaching a focused, robust evaluative conclusion.



© www.julianking.co.nz

Evaluative reasoning

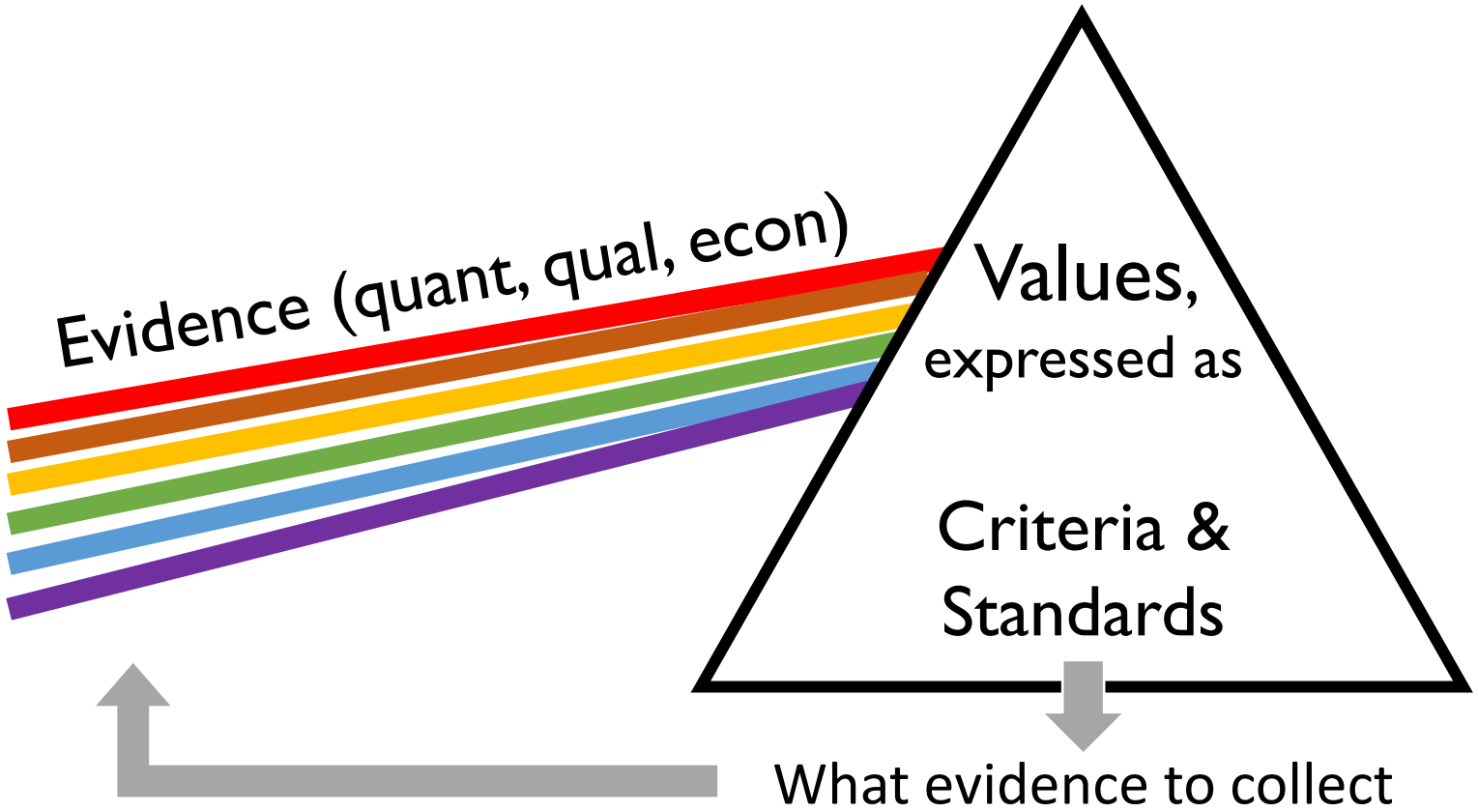
It starts by working with stakeholders, rights-holders and end-users, to define values. Values are expressed as criteria (aspects of VFM) and standards (levels of VFM). Together, criteria and standards provide a statement of what matters, and what good looks like.



© www.julianking.co.nz

Evaluative reasoning

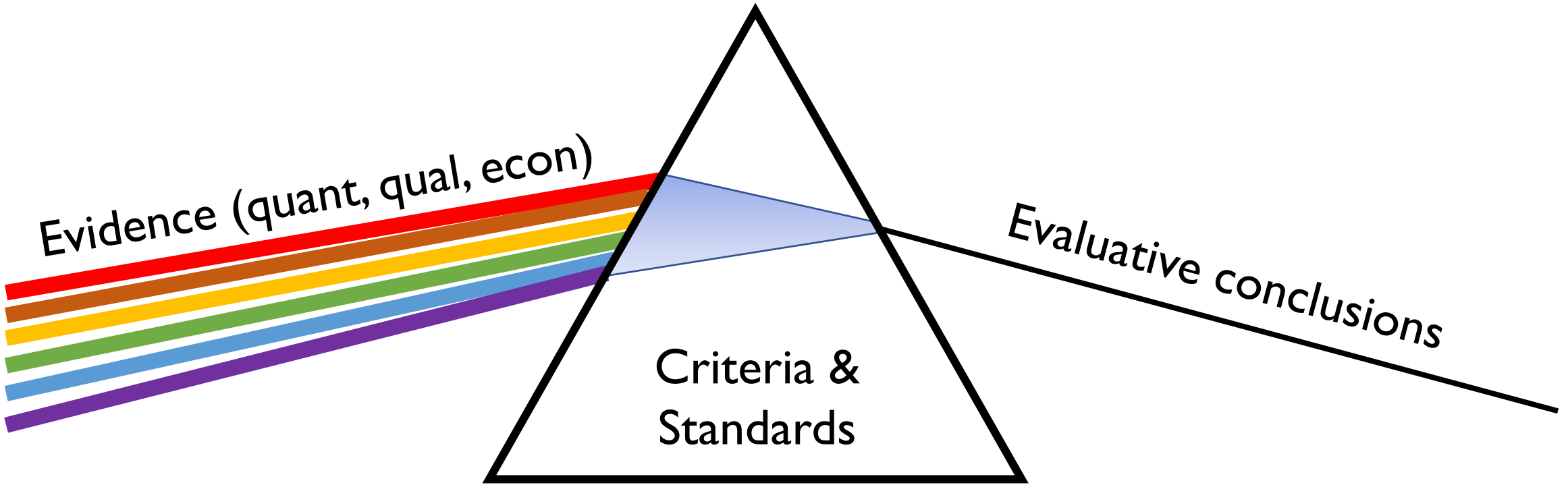
Once we are clear about what matters and what good looks like, we know what evidence we need to collect and analyse.



© www.julianking.co.nz

Evaluative reasoning

And, once we've gathered and analysed that evidence, criteria and standards are a guide for interpreting the evidence, so we can provide a clear answer to the VFM question. This approach builds on the work of Michael Scriven, Deborah Fournier, and Jane Davidson.



What evidence to collect

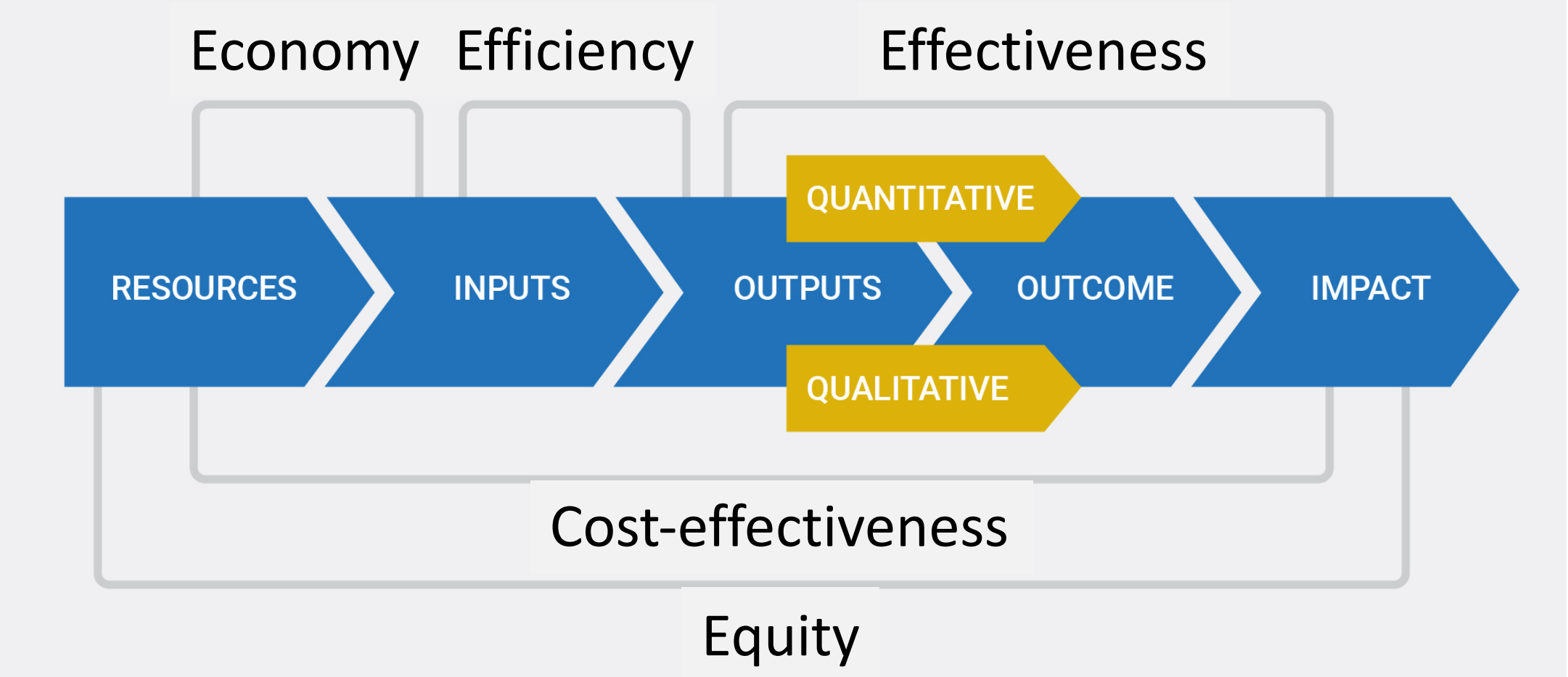
How to interpret the evidence

© www.julianking.co.nz

Criteria = aspects of VFM

Example (DFID, 2011)…

Here's an example of some VFM criteria. These ones come from DFID's approach to VFM (2011). They provide a good starting point, but we need to define them in a way that is more specific to our program and context.



© www.julianking.co.nz

Criteria are specific to the program and context

For example, here's an overview of how we defined VFM criteria for the MUVA program.

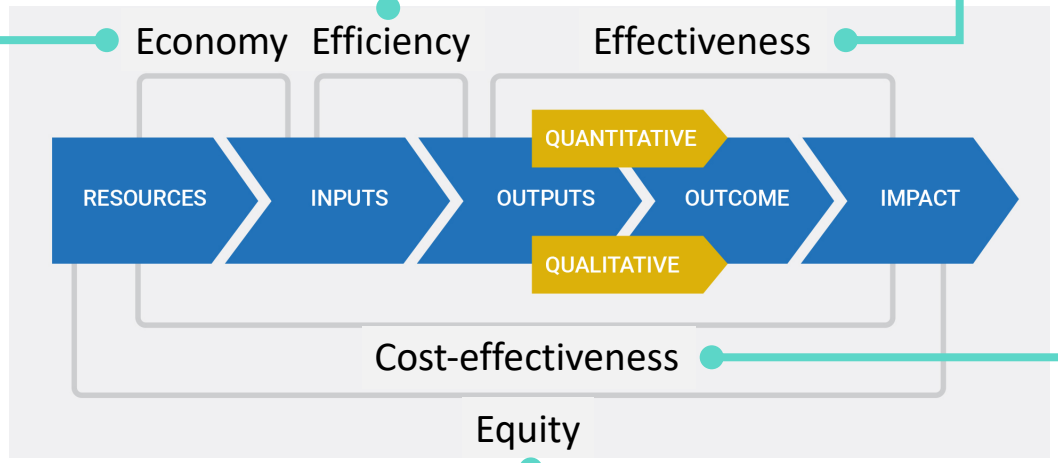
- Good procurement & cost management

Productivity:

- Technical efficiency
- Allocative efficiency
- Dynamic efficiency

Effectiveness as a:

- female economic empowerment program
- learning program
- influencing program



- Equitable design
- Equitable delivery
- Equitable outcomes

- Successful projects, taken to scale, create more value than they consume (economic analysis)

© www.julianking.co.nz

Standards = levels of VFM

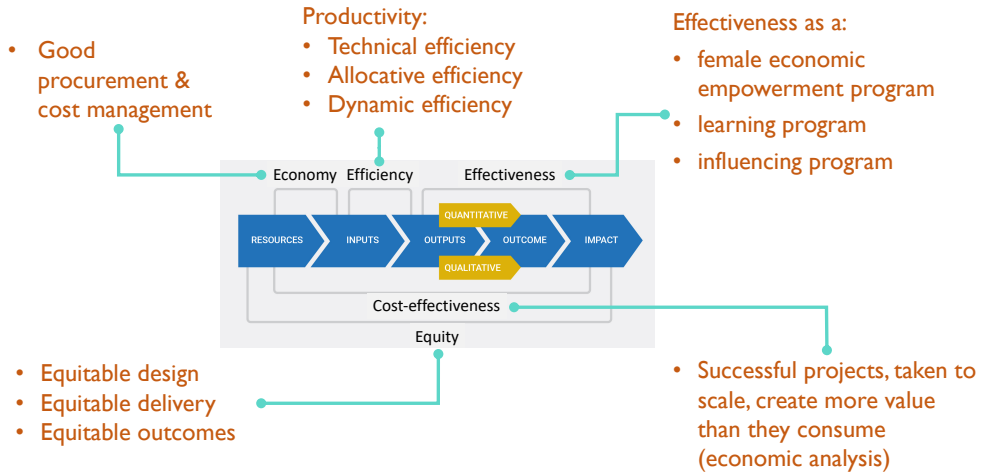
Example (MUVA, 2019)…

Excellent	Meeting or exceeding all reasonable expectations/targets bearing in mind context. Room for incremental improvements.
Good	Generally meeting reasonable expectations/targets, allowing for minor exceptions. Some improvements needed.
Adequate	Not meeting expectations/targets but fulfilling minimum requirements and showing acceptable progress overall. Significant improvements needed.
Poor	Not fulfilling minimum, 'bottom-line' requirements or not showing acceptable progress overall. Urgent improvements needed.

Standards are levels of VFM. This is where we define terms like 'excellent', 'good', and 'adequate'. Once we have defined these terms, we can use them in a precise way in our evaluation. These terms are not superlatives, but carefully-crafted and agreed definitions.

© www.julianking.co.nz

Criteria & standards ➡ what evidence to collect



Excellent	Meeting or exceeding all reasonable expectations/targets bearing in mind context. Room for incremental improvements.
Good	Generally meeting reasonable expectations/targets, allowing for minor exceptions. Some improvements needed.
Adequate	Not meeting expectations/targets but fulfilling minimum requirements and showing acceptable progress overall. Significant improvements needed.
Poor	Not fulfilling minimum, 'bottom-line' requirements or not showing acceptable progress overall. Urgent improvements needed.

➡

Administrative data

Documents review

Outcome evaluation data from pilots (quant & qual)

Stakeholder interviews

Reflection workshops

Cost-benefit analysis of successful approaches at scale

In most cases we find that a mix of quantitative and qualitative evidence is needed to address the criteria and standards. The list above shows some key examples of the evidence we use in MUVA.

Criteria & standards ➡ how to interpret the evidence

Transparent judgements, clear reporting

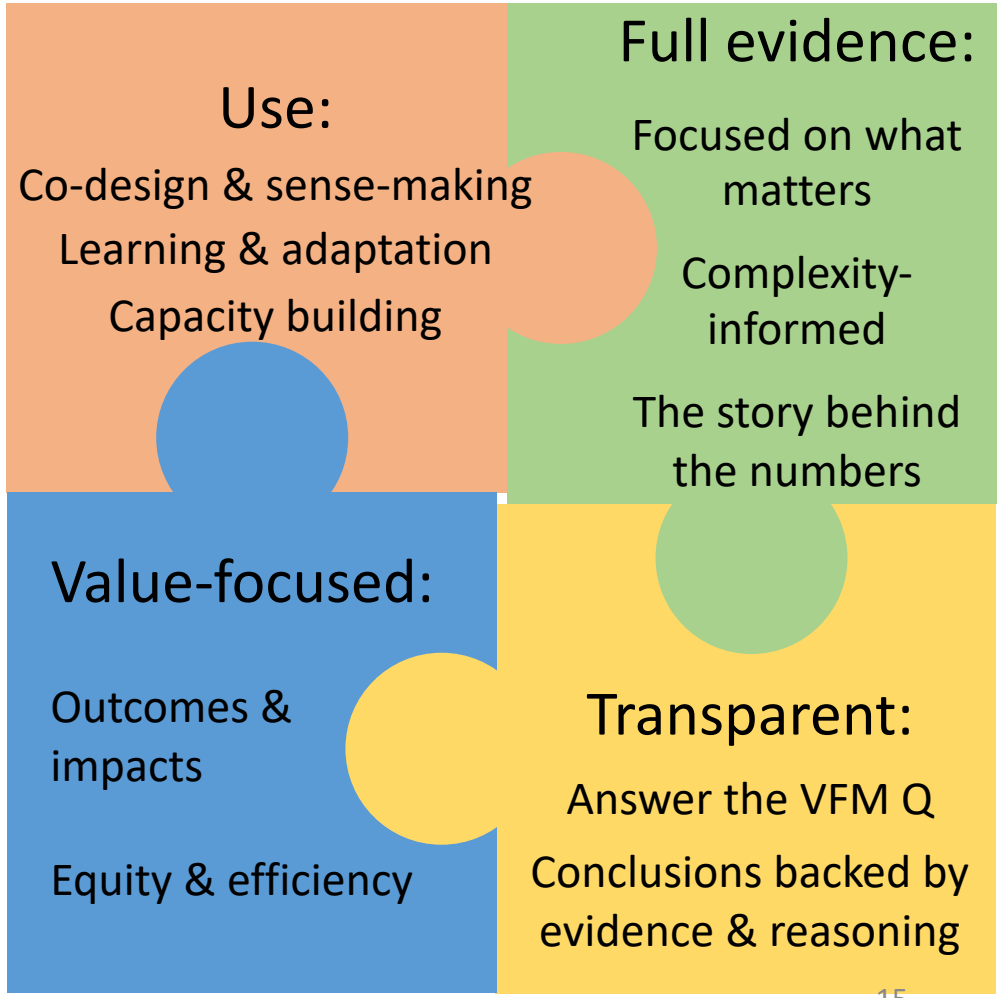
	Judgement	Evidence	Opportunities
Economy	Excellent		
Efficiency	Good		
Effectiveness	etc.		
Cost-effectiveness			
Equity			
VFM overall			

Using criteria and standards helps us get straight to the point and answer the VFM question. This table resembles the format we use in the executive summary of a VFM report. It provides all the information most readers want: the judgement, according to our agreed criteria and standards; the key pieces of evidence that back the judgement; and opportunities to improve.

Aimed at better VFM assessment

from…

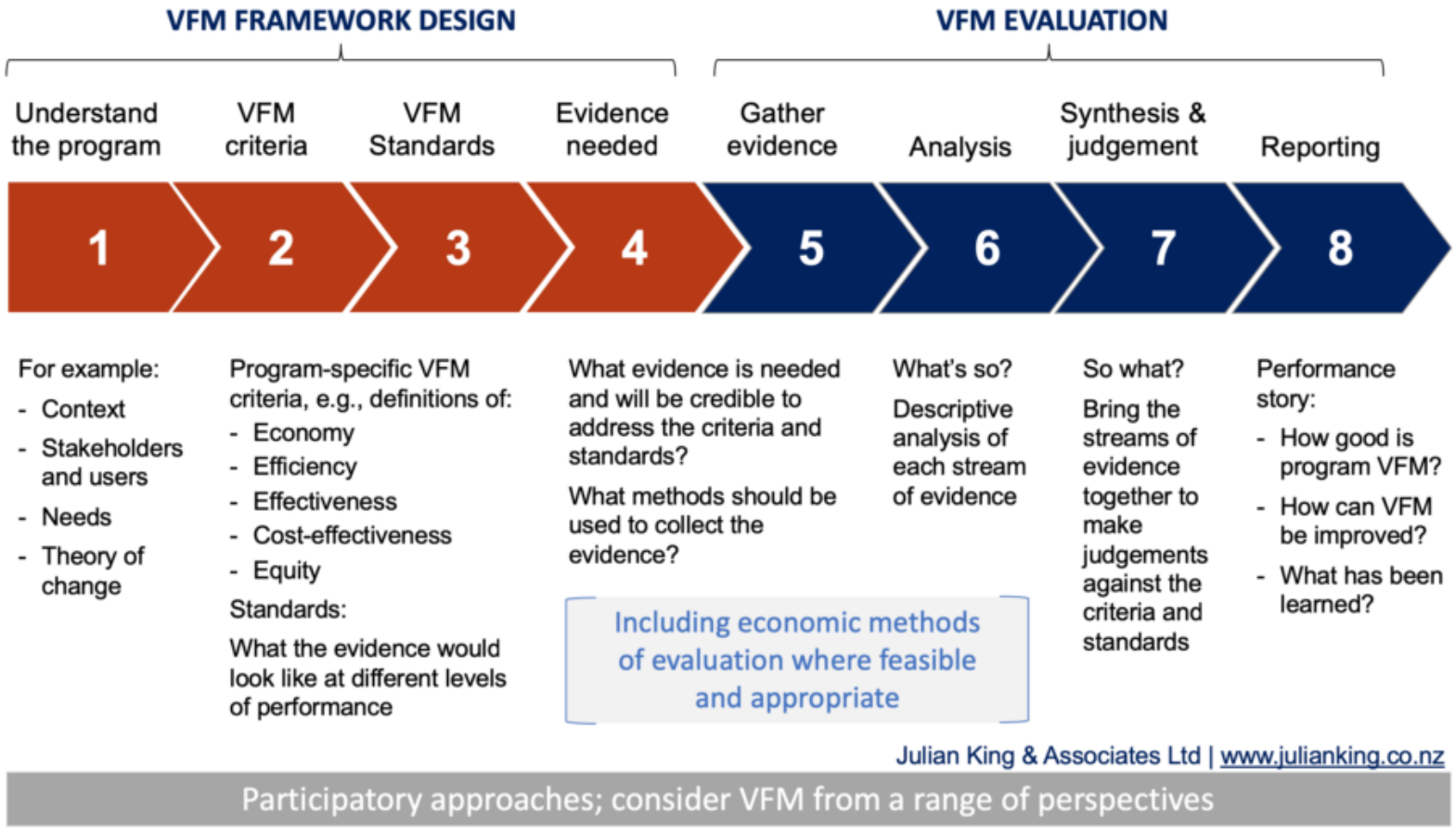
to…



© www.julianking.co.nz

Practical, intuitive process

This approach to evaluative reasoning follows a sequence of eight steps (allowing for some iteration between them). Following these steps helps to ensure the evaluation is aligned with the program design and context, gathers and analyses the right evidence, interprets the evidence on an agreed basis, and answers the VFM question. Each step is an opportunity to involve stakeholders, supporting understanding, ownership, validity, and use.



(King & OPM, 2018)

Why not just use economics?

Isn’t VFM the same as return on investment or bang for bucks?

Well, nearly, but not quite…

VFM means good resource use

The domain of evaluation

The domain of economics

Good resource use is more than just return on investment – it's a matter of context, perspective, values, and trade offs. Thomas Schwandt (2015) described VFM as "the extent to which monetary costs, time, and effort are well used in achieving specific outcomes" (p. 52). DFID (2011) defined VFM as "maximising the impact of each pound spent to improve poor people's lives" and explicitly noted that this includes both efficiency and equity considerations. I have defined VFM as "the merit, worth and significance of resource use" – making VFM the business of both evaluation and economics (King, 2017). VFM is an evaluative question about an economic problem. We should use evaluation and economics together to answer a VFM question (King, 2017).

Costs, consequences and comparisons

Economic evaluation systematically **compares** alternatives

- Cost-effectiveness analysis
- Cost-utility analysis
- Cost-benefit analysis

Economic methods of evaluation include cost-effectiveness analysis, cost-utility analysis and cost-benefit analysis. All of these methods compare the costs and consequences of an intervention, program or policy with alternative uses of the same resources. All three methods value costs in monetary units (e.g. dollars). They differ in the way they measure consequences.

© www.julianking.co.nz

Cost = opportunity cost of foregone alternatives

In economics, every choice we make is at a crossroads. When we decide to invest resources in something, we are forgoing the opportunity to invest in something else. Economic evaluation is concerned with the *opportunity cost* of these foregone alternatives. Costs are not just the money spent in a program, but anything with an opportunity cost. All of the costs are given a monetary (dollar) value, even if no money changes hands. For example, we would count the value of volunteer time, even though it was unpaid.

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) measures costs in money, and consequences in natural or physical units (e.g., number of lives saved, or number of graduates). Results are expressed as a *cost-effectiveness ratio* (e.g. average cost per graduate). Ideally CEA should be comparative, providing an *incremental cost-effectiveness ratio* – the *additional* cost of an intervention, compared to its next-best alternative, divided by the *additional* effects it delivers. Typically CEA is used when comparing two or more ways of achieving the same objective, and the outcome can be adequately measured with one indicator.

Incremental Cost-Effectiveness Ratio:

ICER (B vs A) = ($125,000 - $100,000) / (150 – 100) = extra $500 per additional graduate

© www.julianking.co.nz

Cost-utility analysis

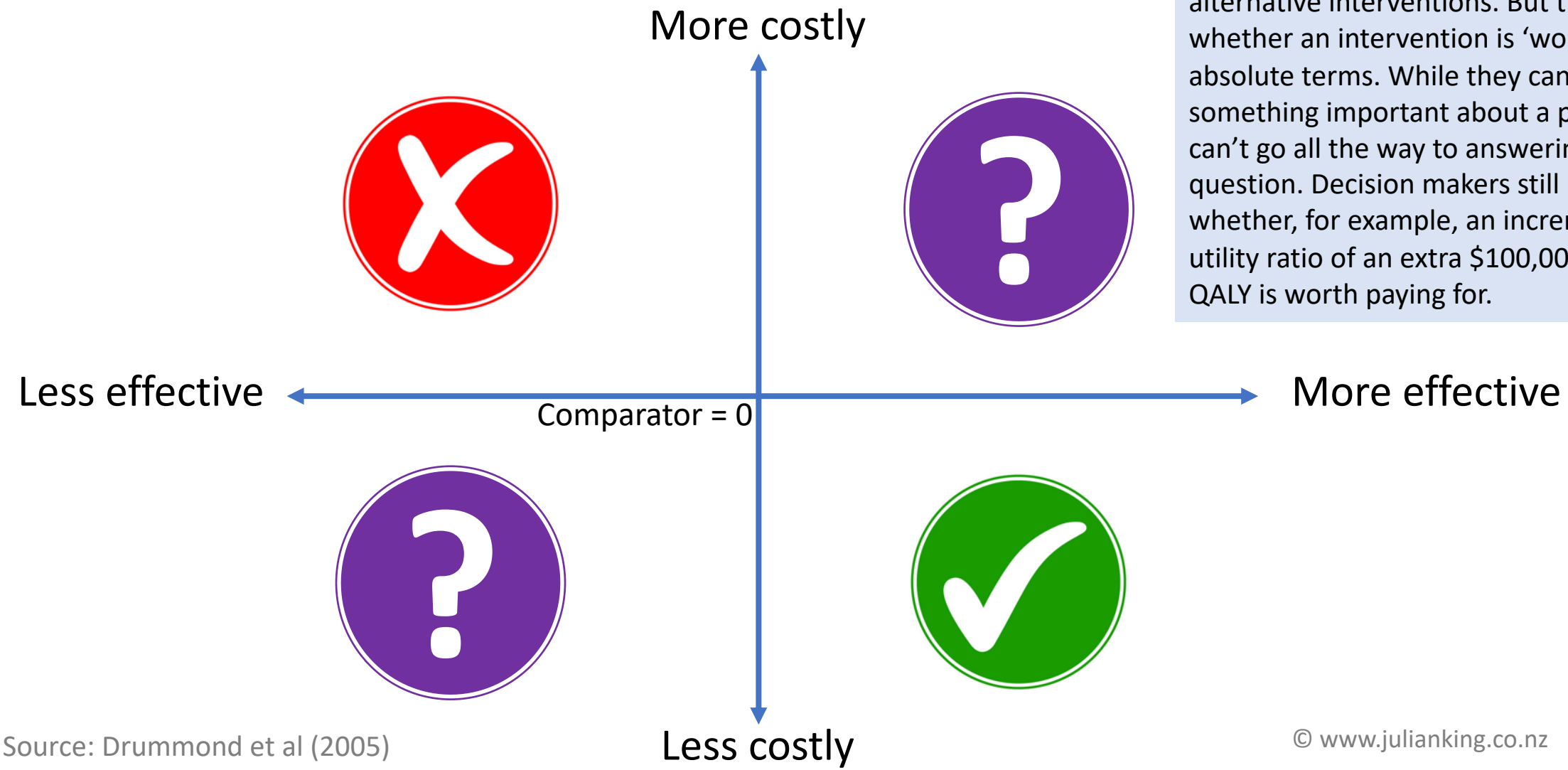
Cost-utility analysis (CUA) is quite similar to CEA, but it includes more information on the consequences side of the equation: the value of the consequences to people. Consequences are measured by an indicator of utility such as quality-adjusted life years (QALY), or disability-adjusted life years (DALY). The output of the analysis is called a cost-utility ratio (average cost per unit of utility) or incremental cost-utility ratio (cost-utility difference of A compared to B). Typical use of CUA is comparing two or more health interventions.

Incremental Cost-Utility Ratio:

ICUR (B vs A) = ($125,000 - $100,000) / (150 – 100) = extra $500 per additional QALY

© www.julianking.co.nz

The cost-effectiveness/cost-utility plane



CEA and CUA tell us about *relative* efficiency of alternative interventions. But they can't tell us whether an intervention is 'worth it' in absolute terms. While they can reveal something important about a program, they can't go all the way to answering a VFM question. Decision makers still need to judge whether, for example, an incremental cost-utility ratio of an extra $100,000 per additional QALY is worth paying for.

Source: Drummond et al (2005)

© www.julianking.co.nz

Cost-benefit analysis

value created → +$

value consumed → -$

One number:
Net Present Value (NPV)

$$NPV_i = \sum_{t=1}^{n} \frac{b_i(t) - c_i(t)}{(1+r)^{t-1}}$$

Cost-benefit analysis (CBA) and related methods like social return on investment (SROI) measure both costs and consequences in monetary units. This has the advantage that costs and consequences can be reconciled into a single indicator, such as net present value (essentially, benefits minus costs), or benefit cost ratio (benefits divided by costs). CBA is comprehensive, at least in principle, because anything can be valued monetarily. Some costs and benefits have an actual monetary value (e.g. the money used to pay for the program; downstream savings; income earned). Some other costs and benefits are bought and sold in real markets (e.g. market wage data can be used to estimate the value of employment outcomes). If real markets don't exist (e.g., what's the value of increased self-esteem or hope?) we can set up experiments using pretend markets to find out what people are willing to pay for things.

Analyst judgements in CBA

Just like any evaluation, economic evaluation is a judgement-oriented practice. Good economic evaluation is transparent about the judgements that have been made. The results of an economic evaluation depend on decisions about what to include in the study, such as...

- **Perspective** – costs and benefits to whom
- **Scope** – which costs and benefits to include
- **Time horizon** – and assumptions about future value
- **Discount rate** – opportunity cost, intergenerational equity
- **Monetization methods** – affects valuations
- **Sensitivity analysis** – which variables, over what ranges

Who, or what, gains or is marginalized by these analytic choices?

(to paraphrase Bob Williams)

Strengths of CBA

Economic methods of evaluation can enhance the validity of evaluation. Two programs could have similar outcomes but might have very different costs. If we don't include costs in evaluation, we would never know that one program is more cost-effective than the other.

CBA has a distinct strength in measuring benefits and costs in the same units, so that they can be reconciled in a single indicator. CBA follows a systematic, rational, replicable approach to counting and valuing the costs and benefits. CBA provides an approximate answer to an important question: *Is society better off overall?*

Often when we conduct an economic evaluation, we are looking at future value as well as past value. When we're estimating future value, we are dealing with assumptions and uncertainty. A strength of economic evaluation is the ability to explore the boundaries of future value through sensitivity and scenario analysis.

Sensitivity analysis tests each assumption to see how stable the overall result is to changes in assumptions. To illustrate, you can say things like "a 10% increase in variable X results in a 2% increase in net present value"

Scenario analysis takes a 'what-if' approach, often changing multiple variables at once, for a range of optimistic and pessimistic scenarios, to see how the result is affected. To illustrate, you can say things like "under the most conservative combination of assumptions, net present value is less than zero"

CBA isn't the whole evaluation

There are things CBA can't tell us, which might be important when determining whether a policy or program is good use of resources. We might want to know, for example, not only whether society is better off overall, but who is better or worse off. The most efficient investments might not be the most equitable. If we want our intervention to reach the most disadvantaged, there may be extra costs. We might want our evaluation to make these trade-offs explicit. If an intervention involves understanding and balancing the interests of different groups, we might need methods that can deal with these differences transparently, rather than valuing everything in the same units. Differences in socioeconomic status, political power, and other factors, sometimes need to be made visible in an evaluation. CBA privileges some types of evidence and values over others. Each **strength** of CBA also reveals a **limitation**.

Quant	**Qual**
Efficiency	**Equity**
Utility	**Other value**
Ends	**Means**
Aggregation	**Deliberation**
Consensus	**Difference**
Majority rule	**Minority voice**
Tangible	**Intangible**
Parsimony	**Complexity**

© www.julianking.co.nz

For example, CBA would struggle if…

- **Program design is evolving** in response to changing context, opportunities, learning
- **We assess VFM in early stages of a program**; too soon to measure outcomes
- **Outcomes are intangible**, e.g. improved teacher self-efficacy, re-uniting refugee family
- **We need a rapid and robust judgement about VFM**, based on qualitative evidence or expert stakeholder workshop
- **Values don't aggregate but instead are in tension** or involve **trade-offs** e.g. costs and benefits to different groups with disparities in political power
- **Program doesn't provide a positive return on investment** but instead creates value through **equity, human dignity, fairness, distributive impacts**

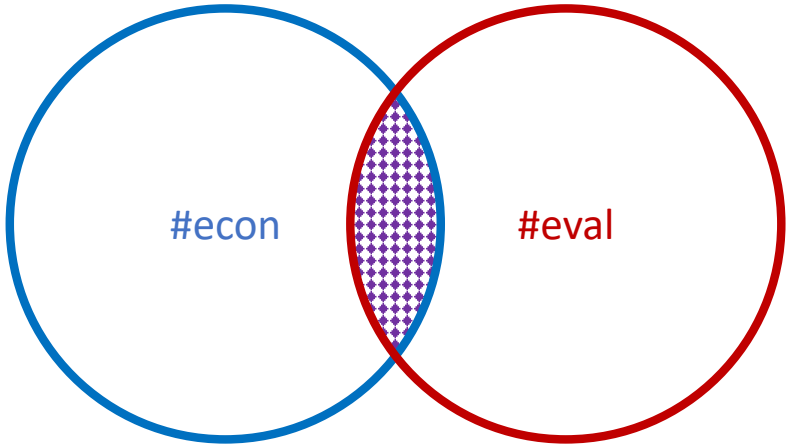
© www.julianking.co.nz

Conclusion:

CBA estimates something important

We should use it more

But not as the whole evaluation



Use economics and evaluation together

© www.julianking.co.nz

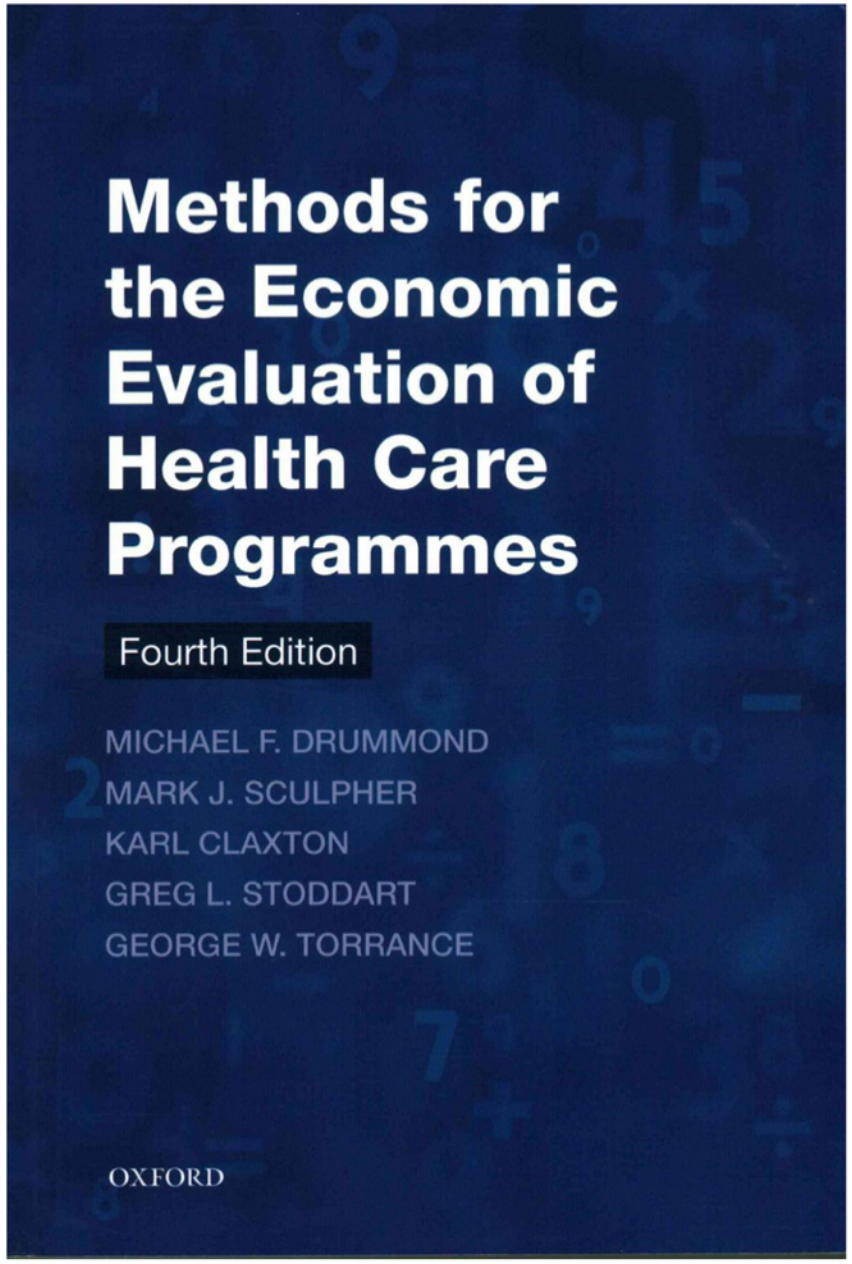
Exercise

Consider a policy, program or project you are familiar with

If you did a CBA:

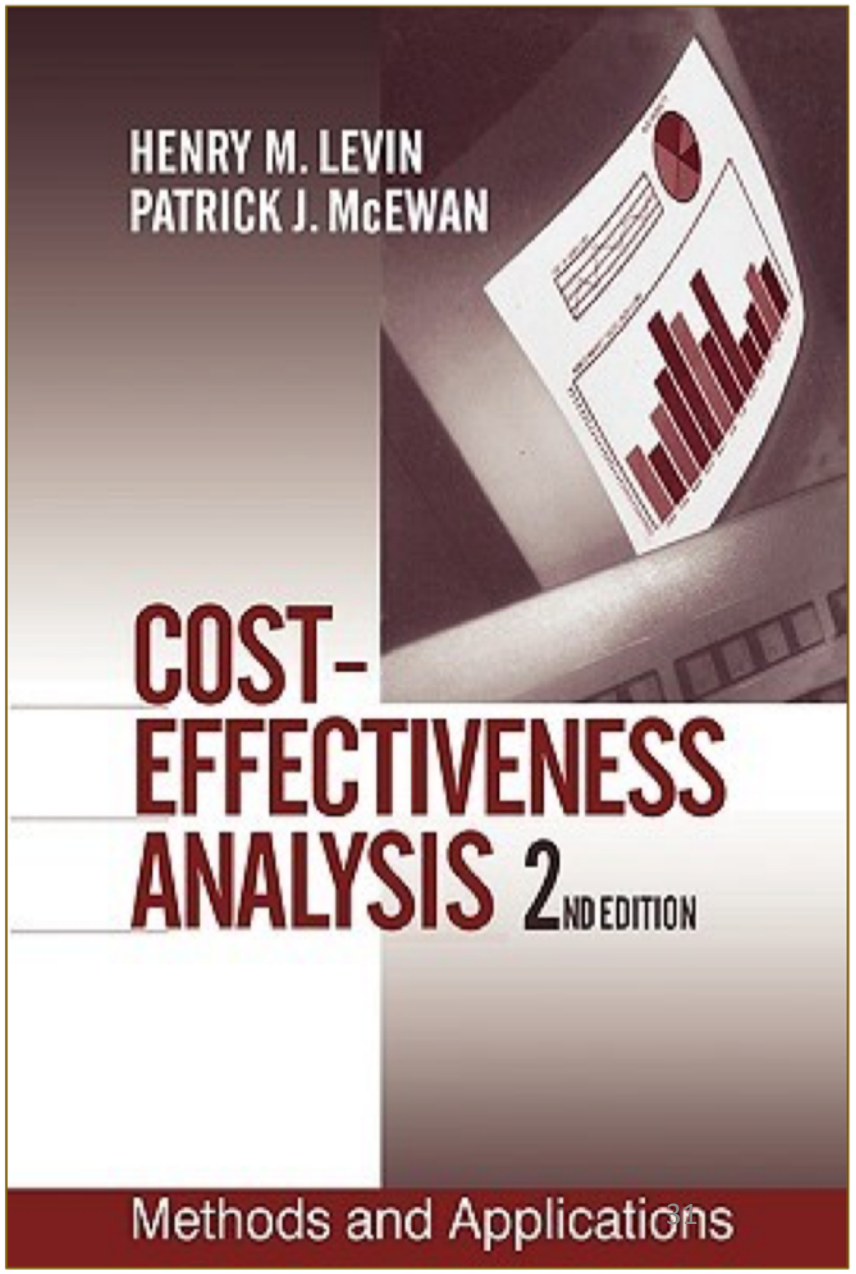
- What costs and consequences would you include?
- Any costs or consequences too hard to include?
- What would the result tell you?
- What would the result **not** tell you, that you might also want to know?
- Who, or what, might be marginalized if you **only** did a CBA?

Economic evaluation resources



Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddard, G. L. (2005). *Methods for the economic evaluation of health care programs.* Oxford, England: Oxford University Press.

Levin, H.M., McEwan, P.J. (2001). *Cost-Effectiveness Analysis*. 2nd Ed. Thousand Oaks: Sage.



Visit my ‘value for investment’ page for publications and blogs:

www.julianking.co.nz/vfi