

# Combining multiple approaches to valuing in the MUVA female economic empowerment program <sup>1</sup>

Julian King<sup>2</sup>

## Abstract

Evaluation and economics each have distinct approaches to valuing. These approaches are traditionally separated by disciplinary boundaries. However, they can and should be combined. Value for money (VFM), in particular, is a shared domain of the two disciplines, because it is an evaluative question about the economic problem of resource use. A theoretical and practical model for combining valuing approaches has been developed through doctoral research. This paper presents and reflects on an example – an international development program where VFM has been assessed using mixed methods (qualitative, quantitative, and economic). Under this approach, evaluative reasoning provides the means for integrating economic values with other criteria and evidence. Deliberation with stakeholders strengthens the valuing process, enhancing validity, credibility and use.

## Introduction

Everybody seems to be talking about value for money (VFM) these days. In the context of constrained resources, competing policy priorities and political pressures to be accountable for stewardship of public funds, decision makers want to know whether policies and programs provide VFM. It's an important question, and it's important that evaluators are equipped to answer it well.

There are some challenges, however. To evaluate VFM, we need a shared understanding of what VFM means. We also need a fit-for-purpose set of approaches for determining VFM under a range of circumstances. I tackled these challenges in my doctoral research, developing and testing a practical model underpinned by theory from evaluation and economics (King, 2019a). In this paper I introduce the MUVA<sup>3</sup> female economic empowerment program in Mozambique as an example to illustrate the application of the model and reflect on its implications for valuing.

Views differ on the meaning of VFM. We understand VFM intuitively as consumers (e.g. deciding what to buy at the supermarket). But when it comes to evaluating a policy or program, one client asks, 'are we getting the right level of outcomes to justify what we're spending?' while another says, 'just reassure us we're not tipping money down the drain'. Similarly, academic and governmental literature offers a range of definitions linking VFM to the economic concept of efficiency, the evaluative concept of worth, normative goals like equity, and bureaucratic goals like sound procurement, among other things (King, 2019a). These are examples of VFM criteria. Schwandt (2015) defined VFM as "the extent to which

---

<sup>1</sup> King, J. (2019). Combining multiple approaches to valuing in the MUVA female economic empowerment program. *Evaluation Journal of Australasia*, 19(4), 217-225.

<sup>2</sup> Director, Julian King & Associates Limited, Auckland; Honorary Fellow, Centre for Program Evaluation, University of Melbourne; Associate, Oxford Policy Management Ltd. [www.julianking.co.nz](http://www.julianking.co.nz)

<sup>3</sup> MUVA ([www.muvamoz.co.mz](http://www.muvamoz.co.mz)) is the program brand name and is not an acronym.

monetary costs, time and effort are well used in achieving specific outcomes” (p. 52). I have generalised the definition to “the merit, worth and significance of resource use” – explicitly making VFM the business of both evaluation and economics (King, 2017).

Views differ, too, on how VFM should be evaluated. I once heard an esteemed professor of medicine proclaim to a packed room that the only way to really know if something provides VFM is to conduct a randomised controlled trial (to determine whether the intervention causes outcomes) followed by a cost-benefit analysis (CBA; to determine whether the outcomes, valued monetarily, are greater than the costs).<sup>4</sup> Evaluation as a field continues to evolve, and has a proud history of methodological debate (Patton, 2018b) and pushback against rigid views about ‘gold standards’ (Julnes, 2012). Any approach has strengths and limitations and thus evaluation methods and reasoning are contextually situated (Greene, 2005; Schwandt, 2015). My research found that CBA brings unique insights that can enhance evaluation (King, 2015; King, 2017), but privileges some evidence and values over others – e.g., quantitative over qualitative, efficiency over equity, ends over means, consensus over difference, tangible over intangible. CBA takes a valid and useful perspective but one that is too restrictive to provide a stand-alone evaluation of VFM in a social policy or program. I concluded that findings from CBA can and should be combined with other evidence and values, contributing to a wider process of evaluative reasoning (King, 2019a).

How, then, should an evaluator go about using economic methods as part of an evaluation? I developed an approach called ‘Value for Investment’ that offers a stepped process and a set of principles to guide methodological decisions (King, 2017; King, 2019a; King & OPM, 2018; King & Allan, 2018; King & Guimaraes, 2016; Kinnect group & Foundation North, 2016; Oakden & King, 2018). The process hinges on developing agreed program-specific definitions of VFM. These definitions provide a framework to ensure the evaluation is (1) aligned with the program design, (2) collects and analyses needed evidence using appropriate methods, (3) draws sound conclusions, and (4) tells a clear performance story that answers the VFM question. These definitions can be set out in rubrics (Davidson, 2005; King et al., 2013), though this isn’t a requirement. What is mandatory is explicit evaluative reasoning, using values (i.e., what matters to people) as the basis of criteria to make evaluative judgements from evidence.

## Example

In this section I describe the MUVA female economic empowerment program in Mozambique to illustrate the Value for Investment approach. Established in 2016, funded by the UK Department for International Development (DFID), and implemented by Oxford Policy Management Ltd (OPM), MUVA has three overarching objectives: (1) developing and testing approaches to female economic empowerment; (2) generating evidence; and (3) influencing others to scale up successful approaches (King, 2019a). DFID requires annual VFM assessments of MUVA. The MUVA team engaged me as an independent evaluator to develop the VFM framework and lead the annual VFM assessments.

The design and intent of the MUVA program exemplifies some common challenges for VFM assessment in social programs. For instance, MUVA is an innovating program, developing

---

<sup>4</sup> For a primer on economic methods of evaluation, I recommend *Cost-Effectiveness Analysis* by Levin & McEwan (2011). Brief descriptions of economic methods are provided at [www.julianking.co.nz/vfi/econ](http://www.julianking.co.nz/vfi/econ)

new approaches in short cycles of testing and reflection. Standard indicators of efficiency, like cost per output and cost per outcome, which are averaged over a defined time period, would be a poor fit with the complex and dynamic nature of the program (King & Guimaraes, 2016). Moreover, there are no relevant comparisons or benchmarks available. Also, MUVA explicitly targets a range of outcomes, some of which can be valued monetarily (e.g. increased earnings) and some that would be challenging to monetise (e.g. participants' acquisition of 'soft skills' like negotiation, listening, self-confidence and teamwork).

MUVA is a learning program, with substantial investment in monitoring, evaluation and reflection-based learning to support adaptive management as well as collecting high quality evidence on successful approaches. The value of learning may be derived from failures as well as from successes, and VFM assessment needs to reflect this value. MUVA is an influencing program, and its long-term value is tied to influencing local organisations, stakeholders and social norms. These outcomes will not become fully apparent during the six-year duration of program delivery – so VFM assessment needs to appraise progress toward these longer-term objectives and the potential future value of scalable approaches (King & Guimaraes, 2016).

I facilitated the development of MUVA's VFM framework in 2016, in a participatory manner with the MUVA team (a mix of local and international staff and consultants, led by OPM) and in consultation with DFID. The participatory approach engaged a range of expertise (e.g., specialist knowledge such as local cultural context, gender and development, and adaptive programming) and ensured the framework balanced objectives of accountability, learning and improvement.

As a preliminary step, to ensure a clear and mutual understanding of the program, we reviewed its theory of change along with other background documents. Then, we defined rubrics comprising criteria (aspects of VFM) and standards (levels of VFM) which were aligned with the theory of change. Rubrics describe what the evidence would look like at different levels of performance (Davidson, 2014). At the broadest level, the rubrics reflected DFID's (2011) five criteria of: economy (which, to paraphrase, focuses on sound procurement); efficiency (productivity in the delivery of outputs); effectiveness (achieving objectives); cost-effectiveness (maximising the value of resources used to achieve outcomes and impacts<sup>5</sup>); and equity (benefiting intended target groups).

At a deeper level, the five criteria were defined in program-specific terms. For example, effectiveness was defined for each of MUVA's three objectives (effectiveness as an urban female economic empowerment program; effectiveness as a learning program; effectiveness as an influencing program). Each criterion was further unpacked with subcriteria specifying aspects of performance that would be assessed. For example, effectiveness as an influencing program included: (1) influencing other DFID programs; (2) effective approaches being taken up and implemented by partners; and (3) stakeholders becoming champions and agents of change (King, 2019a).

VFM assessments have been carried out each year since 2016, and are ongoing. Each year, we use the agreed criteria and standards to help determine what sources of evidence will be needed and credible to address the criteria and standards. The evaluation team gathers and

---

<sup>5</sup> DFID's definition of cost-effectiveness refers to the general notion of comparing value created with value consumed, and not the specific economic method of cost-effectiveness analysis.

analyses evidence using a mix of methods – for example, analysis of financial and administrative data, documentary analysis of monitoring and evaluation reports, key informant interviews, reflection workshops, and economic evaluation.

Assessment against the cost-effectiveness criterion includes an economic approach called break-even analysis. This approach enables rapid assessment of the prospect of successful approaches, taken to scale, generating benefits that equal or exceed costs. Break-even analysis follows the general analytic structure of CBA, valuing costs and benefits monetarily, adjusting them for differential timing, and aggregating them (Adler & Posner, 2006). However, instead of calculating a net present value (essentially, benefits minus costs), break-even analysis has the more modest aim of assessing the prospect of benefits exceeding costs within specified time frames. This approach helps address the problem of missing information, where costs are known but future benefits must be estimated and are subject to uncertainty (Sunstein, 2018). The analysis uses cost and outcome data from pilot projects, together with emergent evidence and transparent assumptions about the approaches at scale. Sensitivity analysis is conducted by manipulating uncertain variables to understand their influence on the result and identify the tipping points where benefits equal costs. For example, what effect size must an intervention achieve to break even, and how achievable is that effect size in light of performance to date?

In each annual VFM assessment, evaluative judgements are made through a process of synthesis, using the rubrics to consider all the streams of evidence collectively, assess the evidenced level of performance against each of DFID's five VFM criteria, and evaluate overall VFM against the five criteria together (King & OPM, 2018). Judgements are a matter of deliberation and debate with the MUVA team and, subsequently, DFID reviewers.

The VFM assessment supports not only accountability but also reflection, learning and improvement. For example, the evaluators prompt the MUVA team to reflect on questions such as: What has been learned about the cost and value of investing in MUVA? What were the smallest investments that had the greatest impact? Were there investments that were quite expensive that in hindsight, did not deliver good VFM? If you reduce a project to its essential elements to make it as scalable as possible, what elements could be stripped back while the approach stays effective? If you could add something (an additional investment) that might bring a disproportionate gain in VFM, what would you add? These questions are effective in helping to embed a 'VFM mindset', supporting adaptive programming and the preparation of successful approaches for scale up.

## **Reflection**

Evaluations often benefit from the use of complementary evaluation methods. Reducing value created and consumed to a net value can be highly informative and I, and others, believe evaluators should use CBA more often (Gargani, 2017; Yates, 1996). However, CBA alone often does not provide enough information to get a full picture of a program and its value. For example, CBA takes a consequentialist perspective, valuing outcomes without examining processes. In contrast we needed to assess the value of MUVA against its core objectives of innovating, learning and influencing, in addition to the value of its actual and potential outcomes. Moreover, CBA provides an indicator of efficiency whereas DFID also requires consideration of equity. More broadly, multiple types of evidence and ways of gaining knowledge should inform evaluative judgements about complex social issues (Deane

& Harré, 2016; Greene, 2005; Wehipeihana & McKegg, 2018). For example, qualitative evidence plays a critical role in MUVA's VFM assessments, in describing the value of reflection and learning processes in adaptive programming. I argue CBA should be used in a supporting role to a wider process of evaluative reasoning and should be used in conjunction with other methods in order to appeal to a wider spectrum of evidence and values (King, 2019a).

This view of CBA highlights the importance of evaluative thinking (Vo & Archibald, 2018) in appraising VFM. Irrespective of the approaches and tools used, evaluation is “not first and foremost about methods, but is about making sense of evidence and creating a coherent, logical, and, ultimately, if successful, persuasive argument about what the evidence shows” (Patton, 2018a, p. 18). Understanding evaluation needs and context, recognising strengths and limitations of methods, selecting and defending contextually-appropriate methods, and interpreting findings accordingly are critical to safe and effective evaluation (Griffith & Montrosse-Moorhead, 2014). Applying these principles to economic evaluation might in some cases lead to a decision not to include CBA (King, 2019a).

Evaluative reasoning (Fournier, 1995; Scriven, 1980; 1994; 1995; 2012) is central to making valid judgements about whether a policy or program provides VFM. Evaluative reasoning is the means by which criteria and metrics from economic evaluation can be combined with wider values, and synthesised to reach an evaluative judgement (King, 2017). The process of developing VFM criteria and standards can also be intentionally used to foster stakeholder engagement and participation in evaluation – facilitating situational responsiveness, validity, and evaluation use (Davidson, 2005; Dickinson & Adams, 2017; King et al., 2013; Martens, 2018). Criteria and standards provide a focal point for engaging with evaluation users and stakeholders, facilitating negotiation and explicit agreement about the basis upon which judgements are made and the types of evidence that are needed and valued. Similarly, stakeholders can be involved in the process of reviewing evidence and making evaluative judgements (King, 2019a). Schwandt (2015) distinguishes technocratic from deliberative and other forms of evaluative reasoning. I argue there are strengths in bringing these approaches together (King, 2019a). For example, in the MUVA program, rubrics are used as a framework to guide deliberation with stakeholders.

### **Challenges and Limitations**

One challenge with this approach would be comparing VFM between different programs. Criteria that are contextually determined for each evaluation may limit the extent to which different programs can be compared or ranked according to their VFM. In anticipation of this challenge, King and OPM (2018) recommended that program-specific standards be calibrated to a generic rubric so that the underlying meaning of ‘excellent’, for example, is consistent across VFM assessments. Where a VFM framework is being developed specifically for like-with-like comparison across multiple interventions, criteria and standards can be specified at a suitable level of abstraction to fit this purpose (King, 2019a). In the MUVA program, for example, such a rubric was developed for assessing the effectiveness and scalability of MUVA's approaches to female economic empowerment.

The ability of CBA to synthesise program outcomes and costs, summarising the result in a single indicator, has beguiling simplicity. In contrast, using criteria and standards to make evaluative judgements from diverse evidence requires a more nuanced and discursive

presentation of findings – though these findings still can and should be summarised succinctly. The output of a CBA can have the appearance of values-neutrality, which some may perceive to be more robust or ‘objective’ than an evaluative judgement based on more complex rubrics and mixed methods evidence. However, this perception would be incorrect. CBA is as much a judgement-oriented practice as any other evaluation method, requiring multiple decisions about things like scope, perspective, time horizon, discount rate, and methods of monetary valuation that affect its results. Furthermore CBA is based upon a normative model of efficiency (Kaldor-Hicks efficiency) with a baked-in assertion that any resource allocation that increases aggregate welfare is desirable irrespective of its distributive effects (King, 2019a). That is the mother of all value judgements, but is seldom questioned in CBA.

Conversely, explicit evaluative reasoning guards against individual evaluator subjectivity. Rubrics, developed with stakeholders as in the MUVA example, are inter-subjective – an agreed social construct used by a group of people for an agreed purpose. Just like other inter-subjective constructs like bank accounts and employment contracts, a good rubric reflects a mutually understood set of values and enables important social processes to occur (King, 2019b). While rubrics reduce personal subjectivity, they can still be affected by shared bias (Scriven, 1991). It is therefore important to guard against cultural biases and groupthink by involving an appropriate mix of stakeholders and perspectives (King et al., 2013) as well as relevant evidence and benchmarks where available (King & OPM, 2018).

## Conclusion

The approach to VFM assessment I describe allows evaluators to bring their full evaluation toolkits, and all relevant available evidence, to determine whether policies and programs are worth the resources invested in them. It offers flexibility to support rigorous, nuanced and context-sensitive appraisals of VFM that respond to stakeholder needs and values in a diverse range of situations.

## References

- Adler, M.D., & Posner, E.A. (2006). *New Foundations of Cost-Benefit Analysis*. Cambridge, Mass: Harvard University Press.
- Davidson, E.J. (2005). *Evaluation Methodology Basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage.
- Davidson, E.J. (2014). *Evaluative Reasoning*. Methodological Briefs: Impact Evaluation 4. Florence: UNICEF Office of Research.
- Deane, K.L., & Harré, N. (2016). Developing a thoughtful approach to evaluation: Values-driven guidelines for novice evaluators. *Evaluation Matters—He Take Tō Te Aromatawai*, 2, 53-78.
- Department for International Development. (2011). *DFID's Approach to Value for Money (VfM)*. London, England: DFID, UK Government.
- Dickinson, P., Adams, J. (2017). Values in evaluation – the use of rubrics. *Evaluation and Program Planning*, 65, 113-116.

- 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.
- Gargani, J. (2017). The leap from ROI to SROI: farther than expected? *Evaluation and Program Planning*, 64, 116-126.
- Greene, J.C. (2005). The generative potential of mixed methods inquiry. *Westminster Studies in Education*, 28(2), 207-211.
- Griffith, J.C., & Montrosse-Moorhead, B. (2014). The value in validity. 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, 17-30.
- House, E.R., Howe, K.R. (1999). *Values in Evaluation and Social Research*. Thousand Oaks: Sage.
- Julnes, G. (2012). Managing valuation. In G. Julnes (Ed). *Promoting Valuation in the Public Interest: Informing Policies for Judging Value in Evaluation. New Directions for Evaluation*, 133, pp 3-15.
- King, J. (2015). Use of cost-benefit analysis in evaluation. Letter to the editor. *Evaluation Journal of Australasia*, 15(3), 37-41.
- King, J. (2017). Using Economic Methods Evaluatively. *American Journal of Evaluation*, 38(1), 101-113.
- King, J. (2019a). *Evaluation and Value for Money: Development of an approach using explicit evaluative reasoning. (Doctoral dissertation)*. Melbourne, Australia: University of Melbourne.
- King, J. (2019b). Is rubric-based evaluation subjective? Blog post. Retrieved from: <https://www.julianking.co.nz/vfi/subjective/>
- 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. Retrieved from: [https://www.nzcer.org.nz/system/files/journals/evaluation-matters/downloads/Online\\_Articles\\_txt\\_King\\_FA\\_0.pdf](https://www.nzcer.org.nz/system/files/journals/evaluation-matters/downloads/Online_Articles_txt_King_FA_0.pdf)
- King, J., & OPM VfM Working Group. (2018). OPM's approach to assessing VfM: A guide. Oxford, England: Oxford Policy Management Ltd. Retrieved from: <http://www.opml.co.uk/publications/opm's-approach-assessing-value-money>
- 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. Retrieved from: <http://idev.afdb.org/sites/default/files/documents/files/Evaluating%20value%20for%20money%20in%20international%20development-.pdf>
- 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.
- Kinnect group & Foundation North. (2016). *Kia E Te Whakangao: Māori & Pacific Education Initiative Value for investment evaluation report*. Auckland: Foundation North.
- Levin, H.M, & McEwan, P.J. (2011). *Cost-Effectiveness Analysis*. 2<sup>nd</sup> Ed. Thousand Oaks: Sage.
- Martens, K.S.R. (2018). Rubrics in program evaluation. *Evaluation Journal of Australasia*, 18(1), 21-44.
- 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. (2018a). A historical perspective on the evolution of evaluative thinking. In A.T. Vo & T. Archibald (Eds.). *Evaluative Thinking. New Directions for Evaluation*, 158, 11-28.
- Patton, M.Q. (2018b). Evaluation Science. *American Journal of Evaluation*, 39(2), 183-200.
- Schwandt, T. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Redwood City: Stanford University Press.
- Scriven, M. (1980). *The logic of evaluation*. Thousand Oaks, CA: Sage.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Scriven, M. (1994). The Final Synthesis. *Evaluation Practice*, 15(3), 367-382.
- Scriven, M. (1995). The Logic of Evaluation and Evaluation Practice. *New Directions For Evaluation*, 68, 49-70.
- Scriven, M. (2012). The logic of valuing. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New directions for evaluation*, 133, 17-28.
- Sunstein, C.R. (2018). *The Cost-Benefit Revolution*. Cambridge, MA: MIT Press.
- Vo, A.T., & Archibald, T. (2018). New directions for evaluative thinking. In A.T. Vo & T. Archibald (Eds.). *Evaluative Thinking. New Directions for Evaluation*, 158, 139-147.
- Wehipeihana, N., & McKegg, K. (2018). Values and culture in evaluative thinking: Insights from Aotearoa New Zealand. In A.T. Vo & T. Archibald (Eds.). *Evaluative Thinking. New Directions for Evaluation*, 158, 93-107.
- Yates, B.T. (1996). *Analyzing costs, procedures, processes, and outcomes in human services*. Applied Social Research Methods Series Volume 42. Thousand Oaks, CA: Sage.

[www.julianking.co.nz](http://www.julianking.co.nz)

December, 2019