

Why Evaluations Fail: The Importance of Good Monitoring

Adam Kessler and Jim Tanburn, August 2014 (links updated Nov. 2021)¹

Contents

1	Introduction.....	2
2	Why should evaluators be interested in monitoring?.....	2
2.1	Good monitoring is essential for effective management.....	2
2.2	Good monitoring is essential for effective evaluation	3
2.3	Some evaluation methodologies incorporate monitoring	3
3	What is the DCED Standard for Results Measurement?	3
4	How does the DCED Standard support evaluation?.....	5
4.1	The DCED Standard promotes clear theories of change	5
4.2	The DCED Standard provides additional data to test the theory of change	5
5	How do evaluations supplement the DCED Standard?	6
5.1	Evaluations are independent.....	6
5.2	Evaluations have more expertise and larger budgets.....	7
5.3	Evaluations can examine broader effects	7
5.4	Evaluations and the DCED Standard are for different audiences.....	7
6	Division of responsibilities between evaluator and programme team	7
7	Key References and further reading.....	9

¹ The authors are indebted to those who commented on early drafts of this paper. In particular, they would like to thank Donna Loveridge and Tim Ruffer.

1 Introduction

A development programme without a strong internal monitoring system often cannot be effectively evaluated. The DCED Standard for Results Measurement is a widely-used monitoring framework, and this document discusses how it relates to external evaluations. Why should evaluators be interested in monitoring systems? How can the DCED Standard support evaluations, and vice versa? Who is responsible for what, and what are the expectations of each? This document expands previous work by the UK Department for International Development (DFID).²

This document is relevant for evaluators, those commissioning evaluations, and practitioners in programmes using the DCED Standard and undergoing an evaluation. It provides a basis for dialogue with the evaluation community; the aims of that dialogue are to identify sources of evaluation expertise available to support programmes using the DCED Standard, and to promote the Standard to programmes needing to improve their monitoring system. We would welcome further discussions on the topic, and invite you to contact us at Admin@Enterprise-Development.org with any questions or comments.

2 Why should evaluators be interested in monitoring?

Evaluation is a 'systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability.'³ There are many types of evaluation; while all forms of evaluation will probably depend to some extent on information generated through monitoring, this paper is particularly relevant to evaluations which seek to understand whether and how programmes achieve their outcomes and impact.

Monitoring is a regular process that systematically collects data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives.⁴ Taken literally, this is a demanding discipline; perhaps as a result, it has been rather neglected by programme managers in the past. For example, Rist and Kusek (2004) suggest that traditional 'implementation monitoring' is primarily concerned with activities, inputs and immediate outputs, and does not examine whether the project led to any significant outcomes.⁵ Monitoring has also sometimes been neglected by evaluators, who stress the need to be detached, and therefore objective.

2.1 Good monitoring is essential for effective management

This limited monitoring function does not provide sufficient information for the management of many modern development programmes, which increasingly aim to catalyse change in complex, rapidly changing systems. In these environments, activities and planned outputs can change as the programme evolves, and it is often difficult to predict the link between these outputs and outcomes. In order to enable experimentation and adaptation, monitoring systems need to support staff to understand what works and why in their intervention. This requires staff to monitor outcomes and examine whether and how programme activities contributed to these outcomes.

² DFID. (2014). [Evaluation and the DCED Standard for Results Measurement](#).

³ OECD-DAC. (2004). Glossary of Key Terms in Evaluation and Results Based Management.

⁴ OECD-DAC. (2004). Glossary of Key Terms in Evaluation and Results Based Management. <http://www.oecd.org/dac/evaluation/18074294.pdf>

⁵ Kusek, J & Rist, R. (2004). Ten Steps to a Results-Based Monitoring and Evaluation System. World Bank.

2.2 Good monitoring is essential for effective evaluation

Weak monitoring systems should be a concern for evaluators as well as for management staff, as development programmes with weak monitoring systems are not likely to be easily evaluable. Riddell (2014) argues that ‘we are still far from having reliable, consistent and robust aid project information to answer the question of whether aid works’⁶, partly because of ‘poor baseline data and the often weak monitoring upon which the assessment of impact is based.’⁷ A 2014 study that examined the effectiveness of Norwegian evaluations noted that only one in five of the evaluations they reviewed made appropriate use of existing monitoring data, and concluded that ‘evaluation is difficult in the absence of an effective monitoring system.’⁸

2.3 Some evaluation methodologies incorporate monitoring

Some within the evaluation community have responded to these concerns with new methodologies and approaches. Most notably, ‘Developmental Evaluation’ seeks to provide real-time feedback to support the development of innovative social initiatives. It recommends a long-term partnership between an evaluator and a programme, where the role of the evaluator is to “elucidate team discussions with evaluative questions, data and logic, and to facilitate data-based assessments and decision-making in the unfolding and developmental processes of innovation”.⁹

Similarly, Pritchett, Samji and Hammer (2011) call for development programmes to complement monitoring and evaluation with ‘experiential learning’. Experiential learning is “the process of disaggregating and analysing existing data to draw intermediate lessons that can then be fed back into project design, over the course of the project cycle.”¹⁰

These and other approaches¹¹ share a belief that evaluative thinking and expertise should be embedded within experimental programmes, to allow them to gather information about what is and isn’t working and why, and use this information to develop and improve interventions in real time. The DCED is keen to hear from anyone engaged in this type of work, to enable mutual learning and exchange.

3 What is the DCED Standard for Results Measurement?

The DCED Standard for Results Measurement is a widely used framework that seeks to enable programme managers to better measure changes, manage implementation, and demonstrate results.¹² It shares a number of features with the methodologies outlined in section 2.3, requiring programme staff

⁶ Riddell, R. (2014). Does Foreign Aid Really Work? An Updated Assessment. Development Policy Centre Discussion Paper No. 33. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2409847

⁷ Riddell, R et al. (2008). Measuring impact: The global and Irish aid programme. OPM.

⁸ Norad. (2014). Can We Demonstrate the Difference that Norwegian Aid Makes? <https://www.enterprise-development.org/wp-content/uploads/Can-We-Demonstrate-the-Difference-that-Norwegian-Aid-Makes.pdf>

⁹ Dozois, E; Langlois, M & Blanchet-Cohen, N. (2010). DE 201: A Practitioner’s Guide to Developmental Evaluation. J.W. McConnell Family Foundation and the International Institute for Child Rights and Development. <http://www.dmeforpeace.org/sites/default/files/Dozios%20et%20al%20Practitioners%20Guide%20to%20Developmental%20Evaluation.pdf>

¹⁰ Pritchett, L; Samji, S & Hammer, J. (2012). It’s All About MeE: Using Structured Experiential Learning (‘e’) to Crawl the Design Space. <https://www.cgdev.org/publication/its-all-about-mee-using-structured-experiential-learning-e-crawl-design-space>

¹¹ For another example, see the Collective Impact Initiative. <https://www.fsg.org/areas-of-focus/collective-impact>

¹² For more information on the DCED Standard, see our website at <https://www.enterprise-development.org/measuring-results-the-dced-standard/>

to conduct activities typically seen as evaluative. In particular, staff must examine whether and how they are achieving their outcomes, and assess attribution to their activities where possible.

The DCED Standard specifies eight elements of a successful monitoring system. Programmes using the DCED Standard begin by developing a results chain to outline their expected results and show how these will be achieved. Practitioners are expected to highlight the assumptions between different levels in the results chain, and provide supporting information to justify them.¹³ Based on the results chain, programmes formulate and monitor indicators, assess attribution and broader changes, and use the results for reporting and programme management. The DCED Standard is implemented internally by programme teams.

The DCED offers an optional audit service, which involves an external, objective assessment of the monitoring system in use in the programme - or for new programmes an assessment of the system in place (but not yet in use). The audit is confidential and will not be made public unless the programme chooses to do so.¹⁴ The audit provides additional reassurance, both to evaluators and wider audiences, that programme monitoring data are of adequate quality.

There is a potential overlap of responsibilities between evaluators and programmes using the Standard, but they remain different tools. The Standard is an internal framework which helps programmes to monitor results, learn and improve. A DCED audit indicates the credibility of self-reported results, checking the monitoring process rather than generating additional data. An evaluation, by contrast, is external and conducted at set points throughout the programme, rather than on an ongoing basis. It can provide additional data to improve accountability and learn more about programme performance.

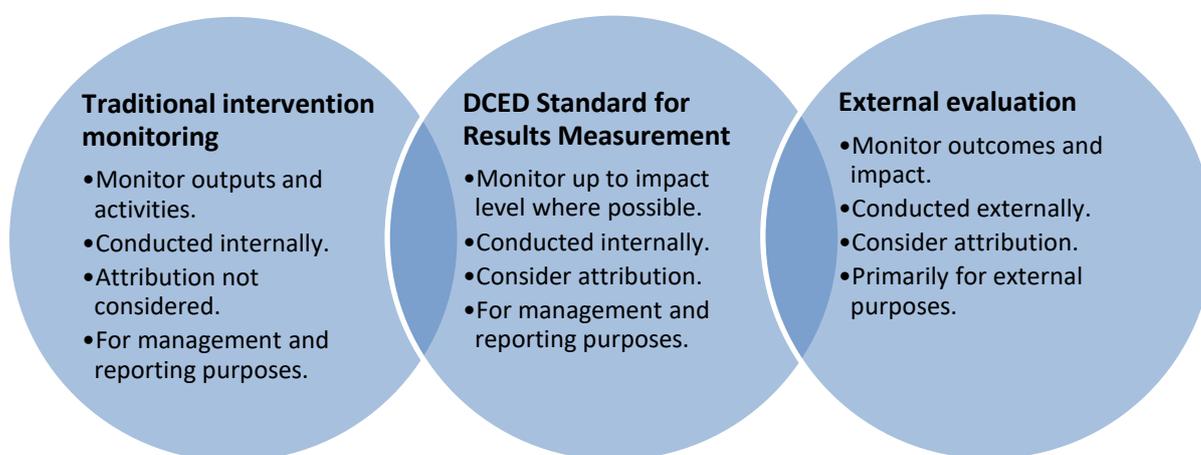


Fig 1: The difference between traditional intervention monitoring, the DCED Standard, and external evaluation.

To date, the DCED Standard has primarily been used with private sector development initiatives, which frequently operate in complex, rapidly changing environments, and have invested significantly in improving the quality of their monitoring systems. The case studies and training materials on the DCED

¹³ A results chain with supporting assumptions can also be referred to as a causal pathway, programme logic, or theory of change. Practitioners using the latter term tend to emphasise the importance of explicit assumptions and an understanding of the context (Vogel, 2012)

¹⁴ For more information, see <https://www.enterprise-development.org/measuring-results-the-dced-standard/auditing-the-monitoring-system/>

website reflect this emphasis.¹⁵ It is, however, a flexible framework which is broadly applicable in almost any type of development programme.

The remainder of this document explores ways in which the DCED Standard and evaluations can work together to improve accountability and opportunities for learning.

4 How does the DCED Standard support evaluation?

Programmes using the DCED Standard are likely to have an improved monitoring system. This can reduce the cost and improve the quality of evaluations by promoting clear theories of change and providing data that the evaluator can use to test them.

4.1 The DCED Standard promotes clear theories of change

Evaluations often seek to understand a programme's theory of change, which 'defines all building blocks required to bring about a given long-term goal'.¹⁶ For example, 3ie's Guide on Impact Evaluation Practice states that 'studies should clearly lay out how it is that the intervention (inputs) is expected to affect final outcomes, and test each link (assumption) from inputs to outcomes (sometimes referred to as the program theory). The evaluation design should incorporate analysis of the causal chain from inputs to impacts'.¹⁷ Understanding and testing the programme theory of change can clarify the evaluation questions, help understand the factors underlying success or failure, and test whether the programme contributed to the measured changes.

If programmes do not have an existing theory of change, evaluators often seek to develop one during the inception stage based on programme documents and discussions with staff and programme partners. This approach, however, has limitations. Programme documents are often out of date and early documents may not reflect how or why the programme theory evolved over time. Staff and partners may not be accustomed to explaining their theory of change, and may not be able to do so in the short time available to an evaluator. Finally, the programme may not have a cohesive theory of change at all – which often renders complex programmes impossible to evaluate.

A programme using the DCED Standard, by contrast, is required to develop results chains for each intervention. This is the foundation of the programme theory of change, showing the different expected changes, outlining assumptions which underlie these steps, and providing supporting evidence to back up these assumptions. The evaluator will doubtless wish to challenge this results chain and may deepen or expand it through the evaluation. The existence of a credible model provides an invaluable starting point which allows the evaluator to start with a clearer, more accurate theory of change.

4.2 The DCED Standard provides additional data to test the theory of change

Evaluators using a theory-based approach typically seek to gather data to test different steps in the programme theory of change, and examine attribution between these steps.

Duvendack and Pasanen (2003) argue that using monitoring data can improve the quality and reduce the cost of evaluations, but highlight three main challenges. Firstly, existing monitoring data are not of

¹⁵ <https://www.enterprise-development.org/measuring-results-the-dced-standard/introduction-to-the-dced-standard/>

¹⁶ <http://www.theoryofchange.org/what-is-theory-of-change>

¹⁷ 3ie guide for grantees (no date) '3ie impact evaluation practice: a guide for grantees', quoted in White, H. (2009). Theory-Based Impact Evaluation: Principles and Practice. 3ie Working Paper 3. https://www.3ieimpact.org/sites/default/files/2017-11/Working_Paper_3.pdf

sufficient quality; secondly they do not provide information about outcomes or impacts; and thirdly they do not look at the causal contribution of the programme.¹⁸

Programmes using the DCED Standard effectively address these three challenges. The DCED Standard requires data on results to be collected according to good practice, and so to be of an appropriate quality. It requires programmes to collect data on outcomes and, where appropriate, impacts, rather than just outputs. Finally, it requires programmes to examine attribution where possible between different links in the results chains. These benefits are enhanced if a programme has had an audit, which checks the quality of the measurement process and provides additional credibility to the self-reported results (depending on the audit score and findings).

The existence of high quality programme data will reduce the need for additional data collection. It will also provide a greater depth and variety of data than the evaluator could have collected, as the programme is able to collect information on a regular basis throughout its lifetime, while the evaluator typically only collects data at set points, such as the beginning and end (or simply the end) of the programme. Of course, the quality of monitoring data will vary in practice, depending on the programme and budget for results measurement. The evaluator will be able to look at the data, examine the paper trail explaining how it was acquired, and make their own judgements on whether it is usable or not.

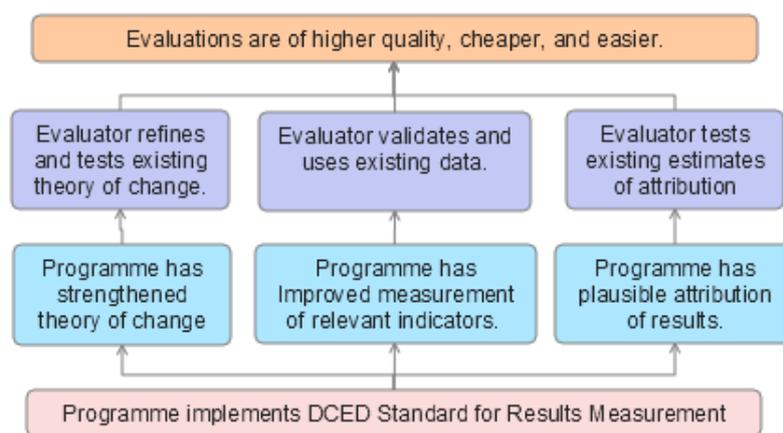


Fig 2: Anticipated benefits for evaluators of programmes using the DCED Standard

5 How do evaluations supplement the DCED Standard?

Evaluations can improve accountability and learning by producing more credible estimates of impact, and going beyond programme monitoring systems to learn more about programme performance. It should be noted that there are many different types of evaluations, and so the following comments will not apply to all of them; however, they illustrate the most common ways in which evaluations can supplement the use of the DCED Standard.

5.1 Evaluations are independent

The DCED Standard for Results Measurement is managed by internal programme teams. Although it may be subject to an external audit, the auditor reports to the programme’s management team. Evaluations, by contrast, are conducted by an independent third party, and the governance structure typically seeks

¹⁸ Duvendack, M & Pasanen, T. (2013). How can monitoring data support impact evaluations? (Blog). Better Evaluation. <https://www.betterevaluation.org/blog/me-data-impact-evaluation>

to ensure the independence and objectivity of their work. This may provide evaluations with more external credibility. It may allow them to challenge assumptions and data quality more robustly than an internal monitoring function can manage, and bring fresh insights and experience.

5.2 Evaluations have more expertise and larger budgets

Evaluators often have more specialised expertise in evaluation methods than programme monitoring teams, and may have larger budgets for data collection. This allows them (in theory) to collect data that go beyond what is achievable by the majority of programmes, gathering data from more sources, taking a more robust approach to attribution and more care to avoid bias. Evaluators consequently may be able to validate (or reject) programme monitoring data, and produce more credible estimates of impact. This deeper and more robust investigation can also yield useful learning which a programme monitoring system is unable to provide.

5.3 Evaluations can examine broader effects

Programme monitoring systems are often somewhat limited in scope. For example, the DCED Standard does not explicitly require programmes to examine cross-cutting issues such as gender, resilience, or climate change (unless they are integral to the theory of change).¹⁹ Moreover, many monitoring systems do not fully examine unintended or negative effects of their work, and few are able to examine longer-term impact beyond the end of the project. Finally, although the DCED Standard recommends that programmes assess the wider impacts of their work on market systems, many programmes find this challenging to do in practice. Evaluations are an excellent opportunity to take a broader view of programme relevance, effectiveness, efficiency, impact, and sustainability²⁰, and contribute further to learning about the broader positive and negative effects of development programmes.

5.4 Evaluations and the DCED Standard are for different audiences

Monitoring systems using the DCED Standard produce information on a regular basis for primarily internal use. The rapid feedback of monitoring information allows for analysis, reflection, and internal learning by programme staff. Evaluations, by contrast, take longer to produce results, and are typically more geared towards external accountability and learning. The difference in emphasis affects the focus of each. For example, as evaluations are typically more geared towards external learning, they may pay more attention to external validity, ensuring that their results are generalizable to different settings.

6 Division of responsibilities between evaluator and programme team

An effective results measurement system should be based on an appropriate working relationship between the evaluator and programme staff, recognising the different contribution that the two can make. This is easier to attain if the evaluator is in place from the beginning of the programme. They can then work with the programme team to allocate data collection and analysis responsibilities, and agree how data can be quality assured and used by both parties. The below table gives an example division of responsibility between programme staff and an external evaluator, drawing from DFID's guidance note

¹⁹ Many programmes are addressing these issues of their own accord. For example, see Markel, E. (2014). Markets Empowering Women: Measuring Results of Women's Economic Empowerment in Private Sector Development, MarketShare Associates for the DCED. https://www.enterprise-development.org/wp-content/uploads/Measuring_Womens_Economic_Empowerment_Guidance.pdf

²⁰ <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

on ‘Evaluation and the DCED Standard for Results Measurement’, and work by Itad on the appropriate division of responsibilities between programme and evaluation staff.²¹

Element of DCED Standard	Role of Programme Staff using the DCED Standard	Role of External Evaluator
Articulate results chain	<ul style="list-style-type: none"> • Develop results chains for each intervention. • Outline assumptions underlying results chain, and evidence for each step. • Review results chains regularly and use to inform programme management 	<ul style="list-style-type: none"> • Review the results chain for quality and clarity, and to understand the programme theory of change • Supplement and validate work done on evidence supporting the results chain by the programme team • Identify and assess the assumptions underpinning the results chain/theory of change • Use the theory of change to identify potential evaluation questions and plan a theory based evaluation approach (if appropriate)
Define indicators	<ul style="list-style-type: none"> • Define indicators for each step in the results chain. • Periodically review and refine indicators. 	<ul style="list-style-type: none"> • Validate the indicators as relevant for evaluation purposes • Define additional indicators that will be required for evaluation purposes and agree responsibilities for data collection
Measure changes in indicators	<ul style="list-style-type: none"> • Collection of monitoring data. • Verification of monitoring data from partners. • Provide evaluator with relevant programme documentation & data 	<ul style="list-style-type: none"> • Validate baseline data collected by the implementer. • Review annually the quality of reporting data for evaluation purposes. • Collect additional evaluation data.
Estimate attribution	<ul style="list-style-type: none"> • Define attribution methods. • Collect data where needed and estimate attribution. 	<ul style="list-style-type: none"> • Review the approach taken and consider the extent to which it can be used for the evaluation along with other evaluative analysis and data gathering • Conduct additional data collection and analysis where necessary.
Capture wider changes in the system or market.	<ul style="list-style-type: none"> • Define expected changes and collect additional data on wider changes in the system or market. 	<ul style="list-style-type: none"> • Review the approach taken and consider the extent to which it can be used for the evaluation along with other evaluative analysis and data gathering
Track programme costs	<ul style="list-style-type: none"> • Track programme costs effectively. 	<ul style="list-style-type: none"> • Validate cost data and use for VFM evaluations
Report results	<ul style="list-style-type: none"> • Produce reports on results. 	<ul style="list-style-type: none"> • Review reports produced and form an independent view / report findings.
Manage system for results measurement	<ul style="list-style-type: none"> • Use results for programme management. • Ensure sufficient human and financial resources available for monitoring. 	<ul style="list-style-type: none"> • Review and advise on improvements to the overall results measurement system

²¹ See DFID. (2014). [Evaluation and the DCED Standard for Results Measurement](#). and Ruffer, T. (2014). Monitoring and Evaluation: The Interface. Itad. <https://www.enterprise-development.org/wp-content/uploads/Ruffer-2.pdf>

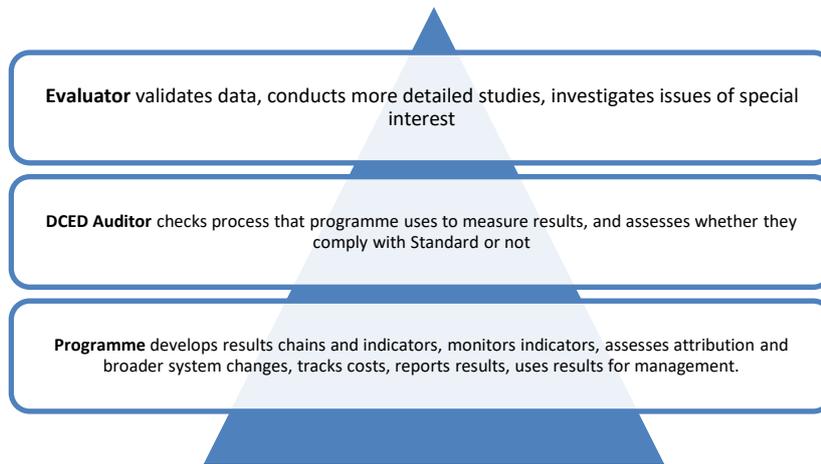


Fig 3: Potential split between programme monitoring systems, audits, and evaluations.

7 Key References and further reading

7.1 Papers:

White, H. (2009). Theory-Based Impact Evaluation: Principles and Practice. 3ie Working Paper 3. https://www.3ieimpact.org/sites/default/files/2017-11/Working_Paper_3.pdf

Dozois, E; Langlois, M & Blanchet-Cohen, N. (2010). DE 201: A Practitioner’s Guide to Developmental Evaluation. J.W. McConnell Family Foundation & International Institute for Child Rights and Development http://www.dmeformpeace.org/sites/default/files/Dozios%20et%20a1_Practitioners%20Guide%20to%20developmental%20Evaluation.pdf

DFID. (2014). Evaluation and the DCED Standard for Results Measurement. https://beamexchange.org/uploads/filer_public/e6/49/e64981c4-9657-4fac-833e-1bfdc3decd68/evaluation_and_the_dced_standard_for_results_measurement.pdf

Kusek, J & Rist, R. (2004). Ten Steps to a Results-Based Monitoring and Evaluation System. World Bank.

Pritchett, L; Samji, S & Hammer, J. (2012). It’s All About MeE: Using Structured Experiential Learning ('e') to Crawl the Design Space. <https://www.cgdev.org/publication/its-all-about-mee-using-structured-experiential-learning-e-crawl-design-space>

Ruffer, T. (2014). Monitoring and Evaluation: The Interface. Itad. <https://www.enterprise-development.org/wp-content/uploads/Ruffer-2.pdf>

Vogel, I. (2012). Review of the Use of ‘Theory of Change’ in International Development. DFID <https://www.gov.uk/government/news/dfid-research-review-of-the-use-of-theory-of-change-in-international-development>

7.2 Blogs:

Bueso-Merriam, J. (2012). The M in M&E: Ugly Duckling or Swan? Inter-American Development Bank <https://blogs.iadb.org/efectividad-desarrollo/en/the-m-in-me-ugly-duckling-or-swan/>