

**EVALUATING
THE IMPACT
OF INTERNATIONAL
DEVELOPMENT
COOPERATION**

A methodological proposal

EVALUATING THE IMPACT OF INTERNATIONAL DEVELOPMENT COOPERATION

A methodological proposal



Acknowledgements

This report was completed thanks to the time and skills of individuals who worked hard to align theory and practice.

We would like to thank:

Stefania Buffa, *Social Value – Human Foundation*

Cinzia Giudici, *Link2007 - COSV*

Cecilia Grieco, *Università Tor Vergata*

Federico Mento, *Social Value -Human Foundation*

Annachiara Moltoni, *Link2007 – ELIS*

Valeria Pecchioni, *Link2007 - CCM*

Giovanni Putoto, *Link2007 - CUAMM*

Maura Viezzoli, *Link2007 – CISP*

Special thanks to the following individuals whose patience and perseverance helped reconcile our different experiences and expectations and led to the completion of the report.

Irene Bengo, *Politecnico di Milano*

Valentina Langella, *Altis - Università Cattolica del Sacro Cuore di Milano*

Margherita Romanelli, *Link2007 - GVC*

Index

SUMMARY	9
1. FRAME OF REFERENCE FOR EXISTING POLICIES AND PRACTICES	17
1.1. The importance of evaluation: current state, beneficiaries and barriers	17
1.2. The challenges of evaluation	19
1.3. The question of methodology.....	21
2. THE ROLE OF AID AGENCIES: AN ANALYSIS OF INTERNATIONAL GOOD PRACTICES	27
2.1. USAID, United States	28
2.2. DFID, United Kingdom	30
2.3. Ministry of Foreign Affairs and Cooperation, the Netherlands	33
2.4. AECID, Spain.....	36
2.5. Agence française de Développement (AFD), France	39
2.6. EU International Cooperation and Development.....	42
3. MEASUREMENT MODELS FOR IMPACT EVALUATION IN DEVELOPMENT COOPERATION	47
3.1. Measurement models and classifications	47

3.2. Synthetic measures	50
3.2.1. Social Return On Investment (SROI)	50
3.2.2. Local Multiplier 3 (LM3)	51
3.2.3. Gamma Model	52
3.3. Process Based Models	52
3.3.1. Process Tracing	52
3.3.2. Outcome Mapping	53
3.3.3. Methodology For Impact Analysis And Assessment (MIAA)	55
3.3.4. Social Impact Assessment (SIA)	57
3.3.5. Impact Navigator	58
3.3.6. Developmental Evaluation	59
3.4. Dashboards and scorecards	60
3.4.1. Fit For Purpose	60
3.4.2. Charity Analysis Framework	62
3.5. Sectoral Frameworks	63
3.5.1. Poverty And Social Impact Assessment- PSIA	63
3.5.2. Health Impact Assessment - HIA	64

4. METHODOLOGICAL PROPOSAL FOR EVALUATION67

4.1. Choice of evaluation model/framework	67
4.2. Review of impact evaluation models in relation to international development cooperation project needs	71
4.3. The approach to evaluation and its integration with Project Cycle Management (PCM)	74
4.4. Guidelines for the measurement of impact generated by development cooperation projects	81
4.5. Step 1 ■ Definizione del contesto di analisi	84
4.6. Step 2 ■ Analysis of objectives and stakeholder engagement	85
4.7. Step 3 ■ Defining the value chain – logical framework	89
4.8. Step 4 ■ Indicator Step	95
4.9. Step 5 ■ Data collection and analysis for impact evaluation	98

4.10. Step 6 ■ Communicating results	102
APPENDIX 1.	104
APPENDIX 2.....	106
BIBLIOGRAPHY	109

Summary

In October 2016, Link2007's General Council commissioned a report on policies that relate to impact evaluation and the evaluation of initiatives and projects in the field of development cooperation. The need to explore the subject in greater depth led to creation of a working group composed of Link2007 network NGO members and external professionals who hold expertise in impact evaluation and in the theory of change methodology.

The end result is a detailed report which includes all the most widely-used methodologies and proposals for future activities that Link2007 network NGO members intend to pilot directly through their interventions. This document summarizes all the work carried out, with a particular focus on tools that are thought to be more applicable to the Italian development cooperation context.

Why impact evaluation?

In line with the mission of NGO members of the LINK2007 network, the impact measurement of actions aimed at eradicating poverty and inequality, securing the fundamental rights of men and women, responding to humanitarian crises, and promoting sustainable development, plays a crucial role in the continual search to bring about positive change.

To this end, starting with the formal adoption of Link2007's 'Document on the journey towards accountability' on 16th October 2013, we have strengthened our commitment to taking measures that place evaluation at the center of our intervention strategies. From this fundamental starting point, we aim to identify and present principles, common guidelines, as well as a set of tools that are useful for impact evaluation. We have brought these together into a document which combines the knowledge acquired and systematizes the wide range of available methodologies.

The following report adopts the OECD definition of impact as “*the positive and negative, intended and unintended, direct and indirect, primary and secondary effects produced by an intervention*”, measured through: a) result indicators (*outcomes*) which embody the direct benefits that an intervention aims to produce and, together with project *outputs* and impact indicators, represent effectiveness; b) long-term indicators which “measure the quality and quantity of the effects produced by the intervention and describe the changes that occur in the lives of individuals and for development at the global, regional and national level, taking into account influential external variables” (OECD, 1991).

Due the inherently complex task of presenting an overarching evaluation framework, we have chosen a more gradual course of action, in which the system of impact evaluation aims firstly, to analyze results in relation to expressed primary needs, and subsequently, to assess long-term effects.

We begin by presenting a summary of the principles that we believe should guide evaluation practice, on the basis of findings from numerous international studies:

Principles for good evaluation practice¹

- Development cooperation agencies should have their own evaluation policy with clearly defined and disseminated guidelines and methodologies, and clear definitions of roles and responsibilities within the evaluative process;
- The evaluation process should be impartial and independent from the policy-making process, in order to achieve greater credibility;
- The evaluation process must be as open as possible with regard to results achieved and made available;
- The evaluation must be used to influence policy-makers;
- The evaluation must involve beneficiaries, partners (local partners in particular) and donors, as well as enabling a learning process that makes the joint achievement of results more efficient and effective;
- The evaluation must form an integral part of program design from the beginning. In order to establish evaluation objectives, it is essential that activity objectives are identified. The evaluation design must define the activities to be evaluated (institutions, sectors, programs and projects) and include information such as context analysis, objectives, resources and expected impact;
- Reporting must be clear and include a description of the activities being evaluated, a description of the methodology adopted, the main findings, the lessons learned and a series of conclusions and recommendations.

¹ Adapted from the OECD/DAC criteria (1991) and UNEG norms and standards (2016);

The final result of our work which is briefly summarized below, is a methodological proposal that can be applied to the activities of development cooperation entities in practice².

Methodological Proposal

Firstly, we chose to review the existing types of model/framework and good practices, in order to broaden the range of measurement models and approaches that organizations can choose from. In this respect, the perception of evaluation as a costly process can be mitigated by the extensive choice of existing frameworks which are highly variegated in their scope, data requirements and analytical perspectives and can guide organizations to make suitable choices based on the characteristics and size of their initiatives/projects and on the availability of resources. In fact, despite the growing importance given to evaluation, the most significant barrier to the development of the practice is the lack of resources in strictly economic terms, as well as time, human resources and the lack of suitable skills, all of which must come from the project itself.

Secondly, this report provides a step-by-step operational guide for organizations that are implementing their chosen model of evaluation in line with Project Cycle Management (PCM), the primary project planning tool in the field of development cooperation. Here, the aim is to integrate the activities and processes of project planning already put in place by organizations, with impact evaluation, namely, to reexamine the elements of the project cycle (process) from the beginning through the magnifying glass of impact measurement.

The choice to present and analyze a wide range of measurement frameworks stems from the conviction that the “one size fits all” model is not easily applied to this context. Models must be chosen following internal organizational reflection and stakeholder engagement. Within the context of international development cooperation, it seems appropriate to standardize evaluation processes, starting from the choice of which methodology to employ, rather than which single evaluation method or method of quantitative/qualitative analysis is most suitable. In fact, there is no single way to conduct evaluation, just as there is no single perspective from which to ‘observe’ the world.

.....

2 Defined according to the legal framework for development cooperation entities, codified in Italian law as ex art 26 legge 125/2014 which includes:

- a. non-governmental organizations (NGO) that specialize in development cooperation and humanitarian aid;
- b. organizzazioni non lucrative di utilità sociale (ONLUS) (Socially-Useful Nonprofit Organizations) that legally operate within international development cooperation and solidarity.
- c. fair trade, ethical finance and microcredit organisations whose statutes have international development cooperation as a primary aim.
- d. migrant community organizations and associations that maintain relationships of cooperation and development support with origin countries or collaborate with entities that fulfil the requirements of the legal article and are active in the State involved;
- e. cooperative and social enterprises, labour and entrepreneur trade unions, foundations, voluntary organisations (law passed on 11th August 1991, n. 266) and associations with social aims (law passed on 7th December 2000, n. 383) in cases where their statutes include development cooperation as an institutional aim;
- f. organisations that have registered offices in Italy and have had United Nations consultative status (ECOSOC) for at least four years.

If the evaluation process that is set in motion proves to be comprehensive and robust and meets measurement needs and objectives, then it will surely include all the rightful components of international development cooperation (i.e. sustainability, empowerment, growth in beneficiary population).

In regard to the choice of which methodology to adopt, there are several factors that can lead an organization to choose a specific model and or combine different measurement methods and analyses. First and foremost, attention must be paid to the beneficiary type (applicant) and to the objectives set, as well as to the underlying values and principles of the actors involved. For example, the use of methods that include participatory techniques is consistent with a specific concept in the field of development cooperation. An additional factor to consider concerns the boundaries of the evaluation, which can be set in relation to an entire organization or to a specific project; in this report, the scope will be the project level.

At the internal organizational level, actors that undertake evaluation activities and processes (decision-makers, *project managers*, project workers) aim to: identify indicators and methods that support internal decision-making, improve the efficiency and effectiveness of activities and community learning, and finally, promote improved social outcomes. Participatory methods allow beneficiary target groups and stakeholders that represent target communities to jointly contribute to project planning and implementation. These stakeholders play an integral role within the delivery phase and are directly involved in the evaluation, alongside managers and project staff. The learning that necessarily follows the evaluation process ensures that impact is strengthened, becoming more consistent and thorough. This leads to a greater responsiveness to the needs of target populations and ensures that the highest number of direct and indirect beneficiaries are reached, both of which are priorities across development cooperation interventions.

Additional guidelines for the selection of an appropriate model relate to the characteristics of each methodology. Approaches which, however sophisticated, are too complex and onerous will not make for an achievable, sustainable and replicable evaluation. Furthermore, each model should ideally be able to **represent the multiplicity of objectives** (social, economic, environmental) sought by a project that regularly produces value for different stakeholders and adopts different definitions of value on the basis of the wide range of cultural settings in which development cooperation interventions take place. The evaluation system must allow for the regular **realignment with social missions**, thus, minimizing the risk of “mission drift”. Moreover, specific project **indicators** must be easily identified and populated, as well as being coherent with the monitoring system that is put in place to preserve the intervention’s multi-subjective nature. Finally, the design of sector-specific models may prove useful when choosing more appropriate indicators to integrate into models that are non sector-specific.

Another important methodological note concerns the concrete opportunity to measure social impacts which by definition can be identified and therefore accurately measured in the long-term. This time frame extends beyond the delivery period and the completion dates of development projects; in order to conduct measurement activities, it is necessary to adopt a monitoring system

that spans 3 to 5 years after the end of the project. This approach cannot be shouldered entirely by the organization delivering the project, but rather, must become a common practice that is promoted and supported, first and foremost, by donor institutions who have earmarked resources accordingly.

To date, the system as a whole does not support organizations to “go back” to their results over time. For this reason, our first objective is to identify and measure *outcomes* as they constitute the best proxy for the expected future impact of an intervention.

Therefore, to summarize, the approach to impact evaluation for development cooperation projects should possess the following characteristics:

- Support the learning and development of the organization;
- Support the empowerment and development of the target community;
- Value the contribution of all actors involved (beneficiaries, partners, institutions and communities);
- Ensure that a positive impact is generated for beneficiaries and that it is used to frame their needs;
- Adopt a method of analysis that is proportional, sustainable and replicable;
- Adopt mixed method quali-quantitative research designs;
- Account for the multiple objectives of an intervention;
- Measure the impact on different stakeholders;
- Account for the alignment between the social mission and results;
- Identify impact indicators that are specific and relevant;
- Carry out the evaluation in line with the monitoring system;
- Communicate impact to investors and donors;
- Communicate impact to beneficiaries;
- Communicate impact to the community and to policy-makers;
- Account for the nature of the intervention (development, reconstruction...).

On the basis of the above considerations and criteria, this document proposes a “systematization” and evaluation of the models/frameworks presented, in order to guide organizations in their choice of appropriate methodologies. The synoptic table contained herein highlights the links between existing evaluation approaches and the aforementioned criteria.

Once the objectives, beneficiaries, resources, time frame and information needs of the evaluation have all been established, the appropriate measurement model and methodology can be selected. The evaluation process proper can then begin and will converge with the various project cycle phases, until it becomes integral to them.

This report is a step-by-step operational guide to the process of evaluation and impact measurement in relation to development cooperation projects. Careful attention is paid to the essential aspects of the practice, while attempting to avoid common mistakes.

Promoting a culture of impact

This report calls for the creation of a culture of impact, whose process logic must become integral to the work of organizations and policy-makers in order to shape the policy-making process.

The existence of cultural barriers at the systemic level continues to limit the use and adoption of impact evaluation practices. First of all, evaluation should not be viewed as the response to an imposed sanction or as a marketing and communication tool. Rather, it should be seen as an activity that can provide important guidance for organizational management, with a view to increasing the effectiveness of interventions. Additionally, it must not provoke fears around the failure to meet unrealistic objectives and the potential of having resources cut if objectives go unmet. Impact evaluation is, first and foremost, a process of internal reflection that improves the knowledge and awareness of organizational activities and ensures a high level of accountability to one’s own stakeholders.

The purpose of impact evaluation lies in the opportunity it provides to understand and learn from mistakes and to grasp experiences and resulting failures or successes that follow the attempts to bring positive and durable changes to the lives of beneficiaries, through increasingly effective and efficient interventions.

In this light, it seems appropriate for the system to progressively build a repository to facilitate the collection of evaluation practices and allow impact evaluation experiences and good practices to be widely adopted and shared. This would greatly increase the level of transparency and improve the effectiveness of projects within the field of international development cooperation.

February 2018

1. Frame of Reference for existing policies and practices

1.1. The importance of evaluation: current state, beneficiaries and barriers

It is by now evident that evaluation plays an increasingly important role within the field of development cooperation. In recent years, the international community has made more consistent and thorough efforts to address aid effectiveness and to consider the role of evaluation processes in improving accountability and providing evidence to allocate resources on the basis of “what works”.

The OECD/DAC guidelines (1991) on evaluation criteria define social impact as the change produced by an aid intervention, whether direct or indirect, intended or unintended, positive or negative. The measurement process can therefore be defined as one in which an organization measures the value of the impact it produces and the stakeholders involved in the process, in an attempt to clarify, measure and report on their ability to produce benefits and highlight the potential for innovation and change embedded within the organization itself (Burdge, 2003; Ashoka). In this respect, the evaluation process can test the assumption that all objectives that produce positive impacts have been met and that those which produce negative impacts have been avoided. The goal is to understand how to direct strategy and decision-making toward the creation of positive social impact for all interested parties (Epstein e Yuthas, 2014).

In the field of development cooperation, the different conferences on Aid effectiveness (Paris 2005, Accra 2008 and Busan 2011) have placed the focus on results by emphasizing the need to consider aid effectiveness. Therefore, in order to better understand the extent to which interventions produce impact, it is necessary to promote processes and methodologies that measure and clarify expected results. An analysis of the impact on development leads to a better understanding of the scope of the change produced by an intervention, as well as to improved and strengthened development programs and policies, including from a cost-effectiveness perspective, where necessary.

From a methodological viewpoint, there is consensus among development actors on the need to adopt an approach that is no longer based solely on the analysis of inputs, activities and outputs, but rather, places greater importance on outcomes, namely, the changes produced by the intervention, and goes beyond the simple process of translating resources into activities. In this sense,

the aim of evaluation is to go beyond the mere understanding of impact in terms of community wellbeing and determine whether and to what extent resources should be allocated in order to meet specific levels of effectiveness and efficiency.

Principles for good evaluation practice³

- Development cooperation agencies should have their own evaluation policy with clearly defined and disseminated guidelines and methodologies, and clear definitions of roles and responsibilities within the evaluative process;
- The evaluation process should be impartial and independent from the policy-making process in order to achieve greater credibility;
- The evaluation process must be as open as possible with regard to the results that are achieved and made available;
- The evaluation process must be used to influence policy-makers;
- Ideally, partnerships should be created which include both beneficiaries and donors in the evaluation;
- The evaluation must form an integral part of program design from the beginning. In order to establish evaluation objectives, it is essential that activity objectives are identified. The evaluation design must establish the activities to be evaluated (institutions, sectors, program and projects) and include information such as context analysis, objectives, resources and expected impact;
- Reporting must be clear and include a profile of the activities being evaluated, as well as a description of the methodology adopted, the main findings, the lessons learned and a series of conclusions and recommendations.

Within the field of development cooperation, there are different actors interested in impact evaluation. Firstly, large institutions and international organizations, national development agencies and researchers have provided a significant contribution to the field, producing and disseminating *guidelines, briefings* and reports to restate the strategic importance of impact evaluation. Secondly, there are several partnerships that involve researchers, institutions and NGOs, which aim to design and disseminate toolkits aimed at practitioners and assess common issues and inaccuracies in evaluation practice. Consequently, there appears to have been a growth in the practice itself: a study carried out by Cameron, Mishra & Brown (2016) on the widespread adoption of impact evaluation points to a steadily upward trend from 2000 onwards. The same study highlights the ways in which, since the 1990s, the growth of evaluation practice has been accompanied by its expansion into new sectors, such as agriculture, education and welfare benefits, spreading beyond the traditional confines of the health sector.

Despite the growing interest and number of actors involved in the practice, there continues to be a gap in evaluation at the global level, which, in 2003, led the OECD to call for “more and better

3 Adapted from the OECD/DAC criteria (1991) and UNEG norms and standards (2016).

evaluation” in order to strengthen the practice within organizations. For this reason, it is useful to turn our attention to a series of challenges that emerge in the practical implementation of the impact evaluation process, which will be further explored in the sections to follow. Among the most significant challenges is the search for a *golden standard* in evaluation methodology. While the value of impact evaluation is well-known – as evidenced by the joint efforts of institutions and practitioners – there are different aspects that continue to inhibit the purposeful and consistent implementation of the practice.

1.2. The challenges of evaluation

The academic and managerial literature that comprehensively analyses the main barriers to impact evaluation practice is broad as it seeks to understand and overcome these barriers, among other objectives.

More generally, it appears as though the greatest limits are presented by the difficulties faced when adopting a less orthodox evaluation perspective, namely, one that includes: a focus on reflection and learning processes, the adoption of flexible methods of analysis that are commensurate with objectives and available resources, and lastly, an analysis of the social value produced for stakeholders. Here, it is important to note that the process of measuring and quantifying social value by linking activities to the qualitative life changes produced for actors involved is perceived to be extremely difficult if not impossible to carry out (Porter et al., 2012). This belief is largely shaped by the skepticism that exists of standardized methodologies which are often seen as “one-size-fits-all” tools that devalue the unique traits of organizations, namely, their social mission, beneficiary characteristics, sector, working context and intervention type, all of which are crucial elements that shape social impact (Ebrahim e Rangan, 2010).

One of the first barriers to emerge in evaluation practice is a cultural one and concerns the actual understanding of what the object of evaluation should be (Epstein e Yuthas, 2014) and the tools available to carry it out (Grieco, 2015). Impact evaluation involves an initial process of theoretical and methodological orientation that begins with *outputs* and ends with *outcomes*, where the former are tangible results that can be directly attributed to a specific activity and are verifiable in the short-term, and the latter refer to a more in-depth change that occurs in the lives of beneficiaries within a broader time frame (and more generally across all groups of stakeholders involved).

In evaluation practice, there appears to be greater emphasis on outputs rather than outcomes. If, on the one hand, the size of outputs can in many cases be used to estimate the value of an outcome that is generally difficult to measure, on the other hand, there is a risk that attention is diverted away from desired project results. The challenge of choosing whether to adopt an output- or outcome-oriented approach is closely connected to the problem of defining the time frame of an impact evaluation. Within the field of de-

velopment cooperation in particular, many evaluations are carried during the project implementation phase and shortly after the end of the intervention. This is a limiting factor since a long-term perspective is generally required to verify that changes have effectively occurred. For this purpose, a good theory of change must be developed in order to establish the time frame within which impact will be achieved.

Cultural limits can also lead to the perception of evaluation as a marketing and communication tool, rather than as a tool for organizational management which can increase the effectiveness of interventions. Funded organizations often fear losing their credibility or resources in the event that their evaluation fails to demonstrate that planned results have been achieved. These fears give rise to “cherry-picking” in which organizations include only the most convenient elements within their evaluations with regard to objectives, stakeholders, indicators, measurement and data analysis methods, in order to emphasize positive results and mask unsatisfactory ones.

Similarly, there is a risk of data being manipulated by organizations that are very keen to demonstrate that targets have been met. In this case, support is required to limit the subjectivity of the organization and ensure that systems which can triangulate data collection and ensure partner engagement exist (Costa e Pesci, 2016). All of this undoubtedly affects the perceived quality of the evaluation itself, calling for an analysis of what, in fact, constitutes a high quality evaluation.

In the field of development cooperation, the evaluative activities of Agencies have often been critiqued for lacking empirical evidence and for their use of approaches that fail to produce valid results (White & Bamberger, 2008). Evaluation runs the risk of being “distorted”, for example, when the information obtained by project beneficiaries is analyzed without the creation of a good control group, or when beneficiaries are self-selected or engaged through available internal organizational databases, which are often limited and do not include individuals from more vulnerable groups.

When conducting evaluation, the most structural barriers concern the poor quality and availability of data (particularly in Developing Countries), which evaluative practice greatly relies on, albeit to different degrees and for different purposes. This type of intangible infrastructure is crucial in order to preserve the reliability of evaluation, improve its comparability and reduce its delivery costs. The lack of available data increases the need to define hypotheses and assumptions, which, in turn, tends to increase the level of approximation and interpretation within evaluation practice (Montesi, 2015). The availability of data becomes an essential precondition, in particular, for the adoption of methodologies that are based on the monetization of impact and require information on the monetary costs and benefits of the social impact of interventions. In these cases, the definition of the equivalent value of the project (for example, in terms of the savings to the state generated by an intervention) could be greatly facilitated by the existence of evidence and socio-economic data on previous public interventions, as well as by cost data on the positive and negative external effects produced at the local and global level (G8 - Social Impact Investment Taskforce, 2014).

The adoption of good evaluation practices is most often limited by the availability of resources in strictly economic terms, as well as by time, human resources and the necessary skills required, all of which must come from the project itself. Agencies such as USAID have earmarked a portion of

their program budget for external evaluation, namely, to commission an independent team. However, the allocation of a fixed amount of funding to evaluation activities can be in contradiction to analytical objectives in situations that require significant resources to ensure a consistent and rigorous implementation, for example, as in the case of experimental methodologies. Conversely, among smaller agencies, there is a tendency to view evaluation as a process which applies only to larger actors whose greater financial resources allow them to make the substantial investments required by the process. In this light, the subject of resource allocation for evaluation activities is key, both in regard to process quality and to the proportionality of the analysis conducted.

1.3. The question of methodology

The debate surrounding the suitability of methodologies clearly deserves separate attention. The choice of which methodology to adopt is not a neutral act in itself as each methodology aims to verify a set of preconditions and priorities on the basis of guiding principles and values, and the choice will greatly affect results. For this reason, its selection must be informed, appropriate and conducted in a transparent manner and there must be an awareness of its validity. In this respect, the primary issue when defining the guidelines of an evaluation is the level of rigor applied and how this concept is defined. The focus on *outcomes* entails a series of challenges, firstly, those that emerge when the change identified and measured is of a primarily qualitative nature. In addition, a long-term perspective requires the actual contribution of an organization to be identified, as several other factors may play a significant role in generating the change.

Shifting the focus onto outcomes requires qualitative changes to be measured, while also ensuring a certain level of reliability and accuracy in the measurement process. In response to this need, greater and often excessive emphasis is placed on strictly quantitative models, namely, experimental and quasi-experimental methodologies⁴ and in particular, on the Randomized Control Trial (RCT)⁵ approach for the construction of the control group. The insistence on methodological rigor stems from the critiques that have been waged against organizations that credit for producing social changes without employing real evidence to back up these claims, other than the use of cases and interviews selected as and when needed.

The advantages of the RCT methodology lie in its ability to attribute the achievement of outcomes directly to a specific intervention, to control factors that could potentially have contributed to cau-

.....

4 Experimental and quasi-experimental methods are statistical designs in which the significance of results is ensured by the rigour of the approach. Quasi-experimental methods have lower reliability due to the fact that the researcher uses pre-existing and pre-constituted samples, with no knowledge of the sampling method employed. As a result, the hypotheses and therefore, the results produced, tend to be weaker. In the case of experimental methods, the researcher controls all the elements of the research design (sample, hypotheses and data).

5 RCT or randomized assignment is a process used to select treatment groups for a program through causal assignment. It ensures statistical equivalence between treatment and control groups, eliminating selection bias.

sing the change and to produce statistically valid data.

For example, in 2016, the World Bank (Gentler et al.) highlighted the need to adopt methodologies that are based on statistically rigorous counterfactual analysis and produce an estimate of the differential change generated by the intervention of an organization, compared to what would have happened in its absence. Similarly, the Centre for Global Development (CGD) reaffirmed the need to adopt rigorous methodologies that allow for the mitigation of risks related to the subjectivity of the researcher in the sampling process for both the experimental and control groups. Nonetheless, it is essential to note that this **methodology is suitable when the intervention under evaluation fulfills set conditions and has certain characteristics. Despite RCT's ability to produce significant results when compared to experimental methodologies, it has several limits that severely hinder its use in certain contexts.**

Firstly, the focus on methodological rigor involves a series of challenges that accompany the adoption of complex statistical models in contexts such as those of Developing Countries, in which the field of reference, the political, social and cultural constraints, as well as the situation on the ground, do not always allow for an appropriate data collection.

This aspect is particularly relevant in the evaluation of interventions that are carried out in the context of humanitarian emergencies (epidemics, earthquakes, droughts, floods, violent conflicts), where it is extremely problematic to use rigorous impact evaluation in order to obtain a systematic understanding of the change generated (Puri et al., 2015). Furthermore, experimental approaches may prove to be unfeasible for interventions that combat illegal practices or behavior, such as, for example, substance misuse, domestic violence, female genital mutilation (FGM), human trafficking etc. The robustness of hypotheses can, in fact, be compromised by the absence of a database and by the challenge of planning and construction a counterfactual. In fact, in such contexts, it is difficult to obtain the data required to assess the beneficiary context prior to the intervention. In the case of projects implemented in certain contexts, in particular, emergency situations, there is often an overlap between the different actors that operate within a given context, making it particularly difficult to identify the actual contribution of an organization. At the same time, in the case of smaller projects which involve low numbers of beneficiaries, the use of experimental techniques can lead project objectives to become misaligned, creating a web of analyses that is paradoxically more onerous than the intervention being implemented. The size of the project being evaluated is also significant as RCT is particularly recommended when a sample size is large enough to ensure a certain level of statistical precision in order to identify program effects. In fact, there are clear specifications concerning the required sample size for an evaluation to be credible and scientifically significant (White et al., 2014). As a result, the increasing complexity of experimental evaluation is also accompanied by significant costs, requiring an adequate level of economic and human resources.

In addition, it is essential to **consider the ethical issues** generated by the very nature of experimental and quasi-experimental methodologies which require there to be clear communication and engagement with all participants. The ethics of this process is particularly important when it comes

to RCT due to the existence of a control group that is not subject to the intervention. The potential disadvantages created for these individuals require that: the randomization techniques used to construct control groups be entirely transparent, the components of both groups be sufficiently informed and finally, that no tensions exist between the treatment and control group. In this sense, it is important to be able to assure the control group that, if the intervention is found to be effective, it will then be replicated for their benefit also. The management of the control group is also a delicate task with regard to the data collection process. In fact, evaluators must consider the fact that they are taking advantage of the availability of individuals who will not benefit from the program in any way. Therefore, they must find compensatory mechanisms to adopt which do not impact on evaluation results (White et al., 2014).

Notwithstanding the methodological issues in relation to evaluation design, experimental and quasi-experimental approaches can provide a “bidimensional” vision of processes of social change, in which, however, the latter are characterized by their complexity as well as their detailed and sometimes contradictory nature, and as a result, are intrinsically “trimensional”. Consider, for example, the unexpected effects of a project, which are also often very relevant, or the negative impacts that may not be identified by an experimental evaluation as they may not be part of the evaluation’s object and will therefore escape measurement. If impact evaluation intends to measure the changes produced in terms of the high-level principles and objectives that an intervention has helped to generate, its tools must be flexible enough to allow for an analysis not only of “what” but also of “how”.

In the face of such limits, the response of practitioners has often been to heavily focus their impact evaluations on the monitoring of qualitative variables and beneficiary perceptions. While this focus does not allow them to produce statistically significant evaluations, it does lead to a more detailed understanding of complex results and their relation to the context. Here, too, there are certain weaknesses evident in practical application and the emerging results:

- The samples used are often too small to be able to generalize the data to the population as a whole;
- The individual and subjective nature of data collection makes it difficult to compare and replicate data.

The **careful consideration of whether to use qualitative or quantitative methods of data collection** can mitigate the risk of presenting an evaluation of the project that is anecdotal or lacks rigor, as well as alleviating the difficulties faced in collecting data that is comparable to other similar interventions. For example, the purely qualitative description of a best practice makes it particularly challenging to replicate or model an intervention on its basis, as well as to objectively verify the results achieved externally. In the same way, a purely quantitative⁶ evaluation has

⁶ For further details on quantitative methods please see Appendix 1.

significant limits, in particular, concerning its inability to comprehend the causal dynamics that underlie the logic of an intervention. This is often due to a superficial analysis of contextual factors which can be difficult to integrate into measurement models based purely on quantitative indicators. If, for example, we were to analyze neonatal mortality in Bangladesh, we would find that the rate is 115 per 1.000 live births for children born in hospital and 67 per 1.000 live births for those born at home. This should lead us to conclude that the mortality rate is twice as high in hospitals. In actual fact, pregnant women who resort to hospitalization have high-risk pregnancies and as such, the statistic takes on a different meaning. The clash between the search for methodological rigor and the need to adopt flexible tools that can be fine-tuned to suit the characteristics of a specific context **has led to the emergence of a new perspective which focuses on the use of so-called mixed methods**⁷. **Here, quantitative analysis is combined with qualitative analysis to both meet the need for scientific reliability and seize the opportunity to delve deeper into each case.** These methods aim to strengthen the validity of results and to understand the process whereby results are produced, while placing them within the context in which they emerge.

An example of this approach is found in the NONIE recommendations – Network of Networks of Impact Evaluation: the adoption of mixed methods allows achieved results that grasp and measure different facets of a complex outcome or impact to be integrated, thus achieving greater depth of analysis compared to if quantitative or qualitative methodologies were adopted singularly. The importance of this approach was reaffirmed by DFID, which points to the opportunity to merge qualitative and quantitative approaches and meet the need for standardized indicators that capture non-material impact and are sensitive to social difference (Stern et al., 2012).

The adoption of mixed methods brings a series of important benefits. Firstly, the focus on the specificities of the context allows for an analysis of the different types of impact that an intervention can have on different stakeholder groups. This is particularly relevant in cases where, for example, the different dimensions of poverty must be measured and where these might include areas that cannot be quantified but that have been identified by beneficiaries themselves (e.g. Dignity, respect, security, power).

The Violence Against Women prevention program launched by Oxfam (America) in El Salvador is an excellent example of the need to adopt mixed methodologies, as well as of its effectiveness. The evaluation encountered several problems; first of all, the challenge of attributing change given the large number of stakeholders and secondly, the impossibility of constructing a credible counterfactual. Part of the challenge lay in the lack of realistic quantitative data that could be used to convey the scale of the problem due to the reluctance of women to report. The use of mixed method methodologies and case comparisons with a focus on intermediary outcomes, allowed for a more in-depth analysis and for data triangulation, which, in turn, led to an understanding of the changes in levels of awareness and in female behavior toward violence.

7 For further examples of mixed methods please see Appendix 2.

In a related study published by InterAction and the Rockefeller Foundation (Bamberger, 2012), mixed methods are defined as the optimal approach to impact evaluation, as they allow evaluators to overcome the limits that emerge when each method is adopted singularly. In fact, in this sense, a purely quantitative analysis may have led to the problem of decontextualization, the “black –box” evaluation approach, while qualitative analysis might have lacked comparability. The study states that the adoption of mixed methods can lead to the following benefits:

- **The opportunity to triangulate results obtained from the evaluation;**
- **The use of results obtained from one type of analysis in order to develop and improve the other;**
- **Complementarity which leads to a deeper and broader understanding of the impact created;**
- **The need to reconcile the potentially diverging results of the two analyses, which necessarily leads to further study and/or to the analytical perspective being reformulated and changed;**
- **The creation of greater value as the use of different methodologies sheds light on different types of value and promotes greater awareness of the evaluation’s overall value.**

Notwithstanding our previous assessment of the different methodologies available, it is the size and characteristics of an intervention that will shape the decision around which methodology to adopt. For example, **the call for greater rigor in impact evaluation processes does not necessarily mean that counterfactual analysis must always be used, but rather, that attention must be paid to ensure that the conditions exist for its correct usage.**

This principle has often been restated. For example, White and Phillips (2012) make an interesting observation when they invite readers to consider the scale of an intervention in order to define the optimal impact evaluation approach. For the authors, if the object of analysis is a “*large-n*” intervention, experimental and quasi-experimental approaches are preferable and RCT is the best option in order to limit the subjectivity of researchers. In contrast, for “*small-n*” interventions, qualitative approaches are preferable in the majority of cases.⁸ In situations where data are only available for a limited number of assigned units, it would be difficult to apply experimental or quasi-experimental methods, as they require the statistical difference between treatment and control groups to be tested.

Another example is from the World Bank, which has identified five questions in order to ascertain

⁸ As specified by the authors, this assumption does not generally apply to all interventions, for example, the quantitative approach is more suited to measure changes at the political-economic level, regardless of the scale of the intervention.

whether quantitative impact evaluation is appropriate (Gertler, 2016):

- Is the program innovative?
- Is the program replicable, can it be scaled-up?
- Is the program strategically relevant?
- Has the program been tested for its effectiveness?
- Is the program influential at the policy level?

In keeping with the above considerations, practitioners have been opposed to the idea of experimental and quasi-experimental methods as widely adoptable solutions and they have identified the conditions in which these methodologies should be avoided and replaced by more qualitative analytical methods such as Process Tracing (see para. 3.3.1). Specifically, these are: when answers are needed quickly, when obtaining precise estimates is not a priority and/or identifying causality is not a primary objective of the evaluation, when the context of reference makes it impossible to clearly identify the attribution effects of outcomes and when the program being evaluated is under development (Shadish et al., 2002).

2. The role of aid agencies: an analysis of international good practices

The present chapter provides an overview of evaluation practices adopted by States which promote development assistance through agencies or governmental bodies. It aims to highlight impact evaluation policies and practices and to present set examples of evaluations in order to examine the procedures and methodologies currently adopted within the field. In this context, the following paragraphs describe the work of development agencies in the United States, United Kingdom, Holland, Spain, France and the European Union.

The chapter points to a growing interest in the subject of impact evaluation and highlights the joint effort that exists to improve the practice and encourage its widespread adoption, in line with international trends. Impact evaluation is most suited to contexts in which the effectiveness and efficiency of a program must be objectively uncovered but also, where there is a need to be accountable and transparent externally. Within their own policies, development agencies identify and state the conditions required for specific evaluations to be carried out successfully, including for impact evaluation. Particular emphasis is placed on programs and interventions that involve significant economic and human resources, for which accountability lies directly with the agency itself. A *program* is defined in line with the European Commission's definition (2006) as a collection of homogenous interventions united by the aim of meeting a clearly-defined global objective. It must specify a time frame and budget and is usually managed and monitored by a specific evaluation or operational unit.

As evident in the course of the chapter and as previously mentioned, although agencies believe that traditional impact evaluation - with its methodological rigor and careful attention to the qualitative dimension - is the tool that best captures the scope of changes produced by an intervention, the practice is not commonly adopted.

In order to plan and justify the adoption of a highly complex evaluation process that requires significant economic resources and efforts, agencies must consider the context, the existence of specific conditions (for example, the availability of credible data) and the questions that the evaluation seeks to answer. The most rigorous impact evaluations are likely to be conducted for innovative or pilot projects, for activities whose assumptions are yet to be tested or for interventions that will potentially be extended or replicated.

2.1. USAID, United States

The United States Agency for International Development (USAID) supports numerous development projects in different global regions and recognizes the importance of evaluation as part of its policy. In response to past criticisms over the lack of quality of its evaluations, the Agency produced a new Evaluation Policy in 2011 and committed to strengthening its evaluation skills, by offering training to internal staff and project partners and creating partnerships with academic experts. This effort is evident in the qualitative improvements made to the Agency's evaluations, thanks to their use of data from diversified analyses in order to validate results, and to their strengthened evaluation practice in general, in line with the new Policy objectives (USAID, 2011).

The Policy document largely contains recommendations and guidelines aimed at the operational units responsible for the implementation of aid programs that are part of American foreign policy. However, it also serves as a model for other agencies and external organizations. In some cases, evaluations are an explicit requirement and provide information and results in order to improve the effectiveness of programs, evaluate the feasibility of expanding or replicating an intervention, and inform the decisions of governments and donors, for example, with regard to resource allocation. Essentially, an evaluation will aim to assess whether a project or activity is meeting its set objectives and how it is doing so, to document good practices and lessons learned and to improve the accountability of organizations. The Policy document identifies the following three situations in which evaluation activities must meet the specific criteria of: **relevance, best method choice, reinforcement of local ownership, transparency and unbiasedness in measurement and reporting:**

- Missions and Washington Operating Units⁹ that manage program funds and develop and implement projects must carry out at least one evaluation per project. This may apply to: the project as a whole¹⁰, a single activity/intervention, or a series of activities/interventions within the projects themselves.
- Missions and Washington Operating Units must **conduct impact evaluation, where possible, for any new and untested approach that may potentially be expanded across US Government foreign assistance or other funding sources, such as a pilot interventions, which must be identified as such during the project development phase.** Impact evaluation must be integrated into project design and if this type of evaluation is found to be unfeasible, a performance evaluation can be carried out with a clear justification for the choice adopted.

⁹ The Operating Units (OUs) implement foreign assistance programmes with the funds assigned for this purpose. OUs include the USAID Mission and the USAID/Dorei/Washington Bureau which use program funds to implement activities with development aims. USAID Missions are responsible for the development and management of the programs, including program and financial monitoring, as well as reporting results.

¹⁰ In the Policy document, the term 'project' refers to a series of planned and implemented interventions identified via a process, which together, lead to specific development results by meeting a related challenge or problem. Therefore the term 'project' does not refer solely to an implementation mechanism as in the case of a 'contract' or 'grant'.

- Each mission must conduct at least one “whole of project” performance evaluation within its Country Development Cooperation Strategy (CDCS) time frame, that covers all its constituent activities and presents the progress made in achieving the project objective.

In order to mitigate a potential conflict of interest, the evaluations required and funded by an evaluation budget must be carried out by an expert that is external to the Agency¹¹ or who has no fiduciary relationship with the implementing partner.

Within the document, importance is given to two particular types of evaluation: *performance evaluation* and *impact evaluation*. The first poses research questions that are typically descriptive, normative and/or relate to cause-effect relationships; it includes a range of evaluation methods and often involves a “before-after” study but generally lacks the rigor of counterfactual analysis, whereas *impact evaluation* measures the change attributed to a specific intervention. The latter is based on a cause-effect model and requires the credibility and rigor of counterfactual analysis in order to control factors that are external to the intervention and may explain the changes observed.

It is essential to design the impact evaluation and select its appropriate methodology in the initial phase of project development to allow evaluators to construct the control group and to collect useful data and key outcome and impact variables well in advance. This ensures that the evaluation can be used to measure change both during and after the program.

By identifying the effects produced by the intervention, impact evaluation helps the Agency, project partners and stakeholders identify the most effective approach in order to meet planned results. This aspect is also fundamental to shape the future development of programs and ensure optimal resource allocation.

In USAID’s vision of impact evaluation, the practice is highly rigorous and therefore adopts experimental and quasi-experimental designs. Nonetheless, seeing as these methods answer specific evaluation questions, they can be adopted in conjunction with other quantitative or qualitative methods, which can shed light on how and why a specific intervention has or has not had its desired impact, as well as making it easier to triangulate evaluation results and understand the context of the intervention.

Finally, USAID states that evaluation should be widely adopted due to its ability to disseminate learning and best practices, as well as to inform and influence decision-makers, with a view to increasing the effectiveness of aid and development programs.

.....

¹¹ Subjects external to the Agency are those from other organizational agencies that are not involved in project implementation or have been hired via A&A (Acquisition and Assistance provides various forms of assistance and support through its own units stationed in the field).



Project Area: Education

Project Promoters and Funders: USAID in partnership with the Ministry of Education of Mozambique (MINED) launched the project *Aprender a Ler* and chose World Education Inc. to develop and implement it.

Project objectives and beneficiaries: The educational project was aimed at second and third grade students in over 1,000 schools in urban and rural contexts within the provinces of Nampula e Zambézia. The objective was to meet two primary outcomes: to improve the quality of reading through education, teacher training and the dissemination of reading materials and to improve the quantity of reading in schools by making school management more effective.

Project evaluation: The evaluation was carried out by International Business & Technical Consultants, Inc. (IBTCI). Through the collection of data on the reading skills of second and third graders, the evaluation objective was to measure the effectiveness of the intervention, its cost-effectiveness and the sustainability of its two levels (improvement in student abilities and increase in school management efficiency). Three groups were selected for analysis (full-treatment, medium treatment, no treatment/control group) using the RCT methodology to ensure greater equivalency between the groups. The chosen methodology allowed evaluators to estimate the impact generated by the project, as well as what might have happened in the project's absence.

Findings: After a full school year, students in treatment schools performed higher both in their reading skills and school attendance, a trend which continued after the Agency's support had ended. The evaluation and the evidence produced illustrated that both school and school management treatment groups constituted the most cost-effective approaches. The fact that the agency incorporated the impact evaluation into the initial program development phase allowed project implementers to quickly alter the program based on emerging findings and recommendations. At the request of the government of Mozambique, the program was extended.

2.2. DFID, United Kingdom

In 2013, the Department for International Development (DFID) released a new Evaluation Policy that replaced the previous 2009 policy (DFID, 2013). The new document is the result of a large internal and external consultation process that DFID held with different stakeholders, including partners, NGOs and foundations and reflects the objective of prioritizing evaluation and making it a key step in the development of interventions. DFID aims to ensure that every year, different types of new and high quality evaluations are commissioned, including impact evaluation, and highlights the need to promote a culture in which evaluation is an essential and routine practice within projects.

This **requires an effort and the necessary investments to be made in order to train staff and strengthen the skills (internal and external) required to conduct evaluation, as well as to carry out new evaluations (including rigorous impact evaluation) that are qualitatively valid.** In order to ensure a high level of quality, it is necessary to resort to independent evaluations to meet the criteria of transparency, utility and ethics.

In DFID's vision, evaluation must consider **whether a planned change has occurred and how the change has impacted on different stakeholder groups.** Consequently, the **choice of evaluation approaches is dictated primarily by factors such as the program context and characteristics, as well as by the evaluation questions.** The expectation is that robust evaluations be adopted and that they **include more than one approach.** Impact evaluation is thought to be particularly appropriate for pilot initiatives where it is important to identify, as early as possible, whether the initiative has been efficient and can therefore be replicated or scaled-up. Furthermore, in cases where there are a large number of beneficiaries, the construction of a control group is advised.

Another DFID principle concerns the suitability of the impact evaluation design in relation to the intervention type and context. In fact, where the RCT method or the use of a control group is inappropriate or unachievable, an alternative method may be used. **The best method is one that allows the data to be analyzed and can establish the validity and replicability of results, as well as reducing subjectivity as much as possible.** DFID uses a *Prospective approach*, namely, **an ex-ante analytical approach which ensures that the measurement process and the necessary data are built into the project from the beginning.**

The Terms of Reference must state the evaluation objective so that the evaluation type and its approach can be identified; furthermore, the feasibility of impact measurement must be assessed by examining the availability and validity of data. Consequently, on the basis of context characteristics and on the intervention type, evaluators must choose whether to adopt a contributive or attributive approach. A qualitative analysis is in any case essential in order to establish how and why there is impact in relation to the intervention context. As such, it is required for every project.

Among the research on impact evaluation commissioned by the department, "Broadening the range of design and methods for impact evaluation" (Stern et al., 2012) and "A guide for Commissioners and Managers" (Stern, 2015) both assess the methodologies that can potentially be adopted to carry out impact evaluation, thus extending the choice beyond traditional approaches. While the first document is highly technical and therefore, aimed primarily at evaluation and research experts, the second highlights the ways in which **traditional approaches to impact evaluation are not suited to the complex and personalized programs of NGOs and civil society organizations and points to the need for a broader range of methods.** The guide emphasizes the importance of the evaluation planning phase as a prerequisite for ensuring the validity, relevance and usability of findings.



Sierra Leone's Youth Reproductive Health Programme (2007 – 2012) - External Evaluation Report (Musa e Gegbe, 2013)

Project Area: Youth and Health care

Project Promoters and Funders: Restless Development Sierra Leone implemented a five-year project with funding from DFID in 135 communities, reaching 146,118 young people. The Restless Development agency works with young people to meet three general objectives for youth: to facilitate the participation of young people in the community (civic participation), to increase employment in order to enable them to contribute to their personal and household income (employment & livelihood) and to improve their sexual and reproductive health knowledge and service access (sexual and reproductive health).

Project objectives and beneficiaries:

The Youth Reproductive Health Program consists of two core elements:

- the rural-based Youth Empowerment Program (YEP) focused on the sexual and reproductive health (SRH) of young people and on in-school and out-of-school life skills education;
- the urban-based Youth Leadership and Advocacy Program (YLAP) focused on building the leadership capacity of YEP alumni and supporting their transition into employment, as well as strengthening their engagement in local and national decision-making processes.

Restless Education continues to adopt its youth-led strategy.

Evaluation: The objective of the evaluation was to carry out a final analysis of the program (2007-2012) that would allow Restless Development and DFID to be accountable to local stakeholders for program results and to estimate the impacts of the intervention in relation to the sexual and reproductive health behavior of young people.

The YRHP Program was evaluated using a quasi-experimental design to estimate the causal impact of the intervention in relation to the sexual and reproductive health behavior of young people. The evaluation design was combined with qualitative and desk research. In addition, a review of relevant documents related to the program was carried out in order to supplement primary data collection. Nonetheless, treatment and control groups were not compared at baseline prior to project implementation and this limited the opportunity to control interfering variables that may have had an effect in the implementation phase.

The treatment group sample size was based on the total population aged 12-28 years in Restless Development's target areas. The control group was selected using a random sampling method seeing as the total number of young people in the target age group was not available. The selection method allowed for an analysis of the effectiveness of the intervention and in fact, it is often employed when the experimental approach is impractical or unethical. While this made the research more feasible, the method presents challenges for researchers in terms of internal validity.

Impact evaluation typically estimates the effectiveness of a program by comparing the outcomes achieved by project participants and non-participants. However, in this case, the main challenge was to find a group that had not participated in the program but that was very similar to the treatment group. In these situations it is difficult to find a group/community that has not been directly or indirectly influenced by the program due to significant cultural closeness and other factors (in this case, listening to radio programs on the subject). Control group results were compared to primary results to present the situation of participants without the intervention.

In order to evaluate the effectiveness of funding use and of learning in relation to program management and development, the data used was obtained through qualitative methods, such as Focus Group Discussion (FGD) to engage young people from both treatment and control groups, key informant interviews and in-depth interviews with young people and other community stakeholders, Restless Development staff and Ministries, Departments and Agencies (MDAs). Finally, closed-ended questionnaires were employed with treatment and control groups, as well as with school teachers.

Evaluation findings (for both YEP and YLAP projects):

- Positive change in behavior among young people in the three parameters of sexual behavior identified (abstinence, condom use and being faithful);
- Young people who took part in the program now use health services and are more likely to identify at least one “modern” method of contraception;
- There is increased knowledge of HIV/AIDS and its transmission routes (reduction in common misconceptions);
- The survey illustrated that thanks to the educational activities offered by the program in schools, young people possess better knowledge of contraceptive methods and use them more frequently. In particular, there has been a reduction in Sexually Transmitted Infections (STI) and in teenage pregnancies thanks to the work of the organization;
- The training activities delivered by the organization which aimed to increase the leadership capacity and participation of young people, have led the community to view these youth in a positive light and as individuals who are able to bring about change, as well as to develop and carry out their own action plan independently (this applies in particular to young ex-volunteers);
- Restless Development has significantly exceeded its targets for young people, in particular, for those who have demonstrated professional skills and in the number of ex-volunteers who have stayed in education or have found employment in the past 12 months;
- Restless Development has managed to successfully engage key government agencies, assisting them in: the development of “youth-friendly” services, with their monitoring and evaluation systems and in the direct engagement of young people;
- With regard to the efficient use of resources, the latest audit report states that 70% of DFID funding (or 72% of total project funding) was used directly for project activities. Additional strategies were put in place in order to strengthen the long-term sustainability of the program;
- Activities that can be replicated in other communities have been identified.

2.3. Ministry of Foreign Affairs and Cooperation, the Netherlands

In the mid-1990s, the Dutch development cooperation strategy changed its approach to development assistance by shifting its focus away from single projects which it considered to be lacking in effectiveness and efficiency due to their fragmented nature and relative absence of coordination,

sustainability and ownership. Instead, it began to directly support the sector and its budget,¹² becoming increasingly aware of the ways in which direct cooperation with ministerial institutions could make development interventions more effective and sustainable, as well as strengthening local capacity.

The Netherlands boast a long tradition in evaluation practice, as recognized by the Agency itself, and evident in internal political debate in the country today which calls for the need to evaluate both the effects and the impact of development interventions. In particular, **evaluation is tied to the policy-making process which requires an analysis of the level of efficiency of programs that are funded and implemented in partnership with local institutions, other donors and organizations.** The evaluation policy guidelines aimed primarily at the department tasked with conducting all evaluations that pertain not only to development, but to foreign policy more broadly (the Policy and Operations Evaluation Department-IOB), focus in particular on the criteria of **relevance, effectiveness, efficiency, impact and sustainability of the results achieved** (Spitz et al., 2013). The work of this department mainly includes impact evaluations on interventions (ranging from programs to various forms of budgetary support) and **policy evaluations (or reviews)** that are summaries of the ongoing studies relating to a specific sector or theme and are tied to, as well as preceded by a series of ex-post impact evaluations.

The same department notes that for certain sectors, it is faster and simpler to find evidence of the impact produced by programs, while in other cases experts suggest that greater attention be paid to certain indicators of quality rather than quantitative indicators in order to measure change. Furthermore, at times, there is a tendency to focus more of quantitative outputs than on quality and impact. One of the criteria at the heart of evaluation is the effectiveness of an intervention or program, in order to determine the causal relationship that connects outputs to desired effects. Nonetheless, internationally and in recent years, greater emphasis has been placed on the search for empirical evidence regarding the effectiveness of government policies. In this context, a single evaluation that describes the relationship between interventions and their effects is deemed to be insufficient rather, it is necessary to identify the net effects of interventions in quantitative terms, where possible. Here, impact evaluation and counterfactual analysis are crucial in order to establish what might have happened in the intervention's absence, together with the level of attribution to the intervention.

Impact evaluation also takes into account the unintended changes that are generated for the intervention target group among other groups, as well as considering how the outcomes achieved have contributed to the achievement of broader, more high-level objectives, such as, for example, poverty reduction.

It is equally important for the evaluation to include both negative and short-term effects (which can be avoided or mitigated) as well as the long-term effects which can have harmful consequences

.....

¹² Budget support is a means or a tool to support development and generally involves the direct transfer of funds to a recipient country government on the part of a donor government.

or even neutralize positive changes. Several qualitative and quantitative methods have been developed in order to measure the effectiveness of interventions. The qualitative approach is seen as providing a plausible explanation for the presence of a specific effect within a given context, however, the results that emerge with this approach are not seen to be generalizable. Conversely, the quantitative approach provides external validity and allows researchers to isolate the effects produced by an intervention within a complex environment, however, it requires a substantial amount of information in order to create a research sample, among other things. The two approaches are complementary and require a sound theoretical framework to be established. The IOB department uses both approaches in its evaluations.

There is a preference for development cooperation activities that favor greater partnership-working between different actors for individual projects, leading to more joined-up, credible, solid and shared planning and evaluation practices.



Turning a right into practice. Ixchen Centre for Women cervical cancer programme in Nicaragua (2005-2009) - Impact evaluation (IOB Evaluation, 2012)

Project area: Sexual and Reproductive Health and Rights.

Project promoters and funders: Within this area, the local NGO 'Ixchen', in partnership with the Ministry of Health, introduced a new approach to turn the right to reproductive health into practice.

Project objectives and beneficiaries: The program focuses on poor women from rural communities in 75 municipalities where the Ministry of Health has limited coverage to detect cervical cancer.

The organization Ixchen provides information, education and awareness-raising activities that aim to increase the knowledge of men and women around the importance of sexual health and the right to demand services that are of a good technical quality and are acceptable (rights-based approach).

Evaluation: The objective of the study is to conduct an impact evaluation of the cervical cancer prevention activities and of the support program (Care Program) implemented by the organization and by the Ministry of Health from 2005-2009. The study focuses on:

- The effect of the intervention in terms of the number of people screened for cervical cancer, compared to the without-intervention situation (net-effect) and the factors that explain this effect;
- The effect on the knowledge and symptoms of cervical cancer and other questions relating to sexual and reproductive health and rights, other than the use of specific services;
- The net effect of the intervention in terms of the number of people correctly treated and the factors that explain these effects;
- Program results in terms of the creation of public-partnership created;
- An analysis of the cost-effectiveness of the intervention;
- The unintended effects of the intervention.

The study was set up as an impact evaluation with certain elements of counterfactual analysis and combined both qualitative and quantitative methods.

The quantitative method includes a survey with 634 users (whose screening results are both positive and negative), an analysis of a database that includes 4.432 users with positive screening results and the results of a cost-effectiveness analysis. Finally, the study also includes a qualitative section composed of 79 interviews with users and their families, as well as non-users, Ixchen staff, the Ministry of Health and 5 case studies.

Findings: The findings confirm that screening and treatment interventions for cervical cancer produced substantial “net effects” for treated users, although knowledge levels remain low. The study calculated that the number of years of life saved through the program was between 6.500-12.500 at a cost per life saved that ranges between US\$ 55 and US\$ 348.

The evaluation processes highlighted the success of the program in encouraging women to look for services that they would not otherwise have sought. Therefore, a positive net effect was identified in the number of women who undergo screening. Conversely, knowledge of cervical cancer remained low (there was little difference between Ixchen users and non-users).

Results indicate that the quantitative component of the impact evaluation did not assess the level of knowledge of non-users, making it difficult to identify the characteristics of the approach used and its effects on levels of knowledge. Cervical cancer screening and support programs require standardized and rigorous management and data collection systems.

2.4. AECID, Spain

The Agencia Española de Cooperación Internacional para el Desarrollo (AECID) has produced a policy on Spanish development cooperation (together with quadrennial Master Plans¹³) in which a veritable evaluation culture is promoted as a system that can contribute to improving the effectiveness and quality of development, as well as facilitating transparency and accountability.

Evaluation should not be viewed solely as a final stage in the management of a project, but rather, should form an integral part of each phase in order to promote the achievement of Spain’s human development objectives, the reduction of poverty and the promotion of human rights. Among its evaluation criteria, the methodological manual produced by the AECID (2013) includes the concept of impact in line with recognized international practices. Importance is given to assessing the changes produced in relation to desired objectives which must be identified in a clear, meaningful and measurable way *and* to the analysis of impact generated in the medium and long-term, alongside any other expected and unexpected changes in relation to the original project plan.

This type of analysis requires a net effect to be established and therefore, for a distinction to be made between what is attributable to the intervention being evaluated and what appears to be the result of other factors.

In evaluation practice, data are gathered using qualitative or quantitative methodologies (surveys, focus groups, interviews, structured observation) in an attempt to balance the two methodologies and maximize participation.

Within Spanish international development cooperation, development NGOs play an increasingly important role, largely due to the fact that their work broadens the sectoral reach of development. Their activities have a longer-term scope and as such, they receive more resources from the Agency.

The important role played by NGOs in development policy, as well as their management of a significant proportion of development funds is evident in the most recent regulations published by the AECID which outline the role of evaluation in the work of NGOs. This is why an interim evaluation is required for all contracts stipulated and a final evaluation for all projects that receive 350.000 euros and above in funding. The regulations state that evaluation must be viewed as an ongoing activity that concerns the entire project cycle, in line with the Agency's methodological manual.

Similarly, the planning phase in which the needs and challenges that must be met by an intervention are established, is seen as very important. Alongside the identification of needs, this phase involves the selection of tools to be adopted in order to evaluate impact and to define whether and to what extent the intervention contributed to improving existing conditions.

The Spanish Impact Evaluation Fund (SIEF) was set up in 2007 following negotiations between the Spanish Government (Foreign Affairs & Cooperation and Finance Ministries) and the World Bank (Human Development Network) and is the clearest example of the commitment that Spanish development cooperation has made to impact measurement. The fund aims to support the impact evaluation of programs aimed at improving human development outcomes and follows the principle that evaluation findings can contribute to the ultimate objective of increasing the effectiveness of development policy.

The calls for funding published require applicants to specify a clear counterfactual evaluation design. In fact, randomized design is the most frequently used technique in impact evaluations funded by SIEF.

The fund has supported the production of a series of policy notes for dissemination; it has organized specific workshops on impact evaluation to increase knowledge and capacity within the sector, and has published the practices gathered in several books, as well as in a textbook, *Impact Evaluation in Practice*, edited in collaboration with the World Bank.



Servicio para la inserción laboral de personas con discapacidad en Ecuador: trabajando por la inclusión (Latas, 2016)

Project area: Disability inclusion

Project partners and implementers: funded by the AECID the ONCE Foundation, implemented by the Confederación Española de Personas con Discapacidad Física y Orgánica (COCEMFE) and by the Federación Nacional de Ecuatorianos con Discapacidad Física (FENEDIF).

Project objectives and beneficiaries: To allow people with disabilities to exercise their right to work, promoting greater equality between men and women in Ecuador.

Evaluation: A joint external evaluation was conducted with the following objectives:

- *evaluation of results:* whether set objectives, results and activities were reached/delivered, highlighting potential impacts in the medium-to long-term which are useful for future activities and to guide the interventions of the various actors involved;
- *impact evaluation:* evaluation of the intervention's effects on the beneficiary population.

Impact evaluation questions:

- what impact (positive/negative) has there been for beneficiaries (entry into work? If yes, on what terms?);
- how have educational activities affected the ability of beneficiaries to improve their family business or to create their own business?

The project evaluation methodology was defined in line with planned objectives, using both quantitative and qualitative tools. For the former, beneficiary data on entry into work was analyzed while the latter involved an assessment of the effects produced by the project and as perceived by users. The quantitative analysis provides evidence of the impact of the employment services for people with disabilities, highlighting the role of the existing legal framework in favoring service take-up. The impact of the project is not only evident from a quantitative point of view in fact from a qualitative point of view the perception of workers was equally positive. This trend emerged in areas that range from entry into work, the subsequent increase in autonomy and the ability to support one's family, to the increase in self-esteem and in the skills acquired through the educational program.

Findings: The main effects perceived by users who entered the workplace are tied to their actual entry into work, their increase in income, the growth in their self-esteem and their strengthened skills. Among the entrepreneurs, the primary evidence of impact was found in: levels of access to education and/or training, technical support to obtain credit, greater spending power, an increase in opportunities for life changes and the recognition as financial creditors.

2.5. Agence française de Développement (AFD), France

Evaluation plays an important role for the French Agency, which, for years, has committed to funding and undertaking evaluation both as an end to itself, as well as for research purposes. In the past years, the Agency has taken part in debates on impact evaluation, participating in various international networks and creating major partnerships with research institutes such as J-Pal with whom it has carried out impact evaluations. The Agency's work in this area is guided by the objective of strengthening and promoting evaluation practice, as well as broadening the range of methodological approaches available in order to make evaluation as effective as possible and adapt it to different contexts.

The experience acquired within the sector thanks to the human and economic resources invested, allows the Agency to broaden its knowledge and skills and to be influential in international debates.

In order to strengthen its evaluation activities, the AFD recently published a Policy document (2013) that is part of the French Official Development Assistance (ODA) and is used by the principal public sector actors in the field of development.¹⁴ The document is addressed to both bodies and Agency staff, as well those they work with: partners, implementers of funded projects, NGOs and donors.

As a technical and financial stakeholder, direct involvement in evaluation processes allows the AFD to better understand and control the complexity of its operations and as a result, to have clearer idea of the scale of results achieved.

Although impact evaluation is considered important and is well-funded, it does not replace other types of evaluation. The Agency's policy identifies four main types of evaluation, based on the assumption that impact evaluation cannot easily be applied across contexts and that evaluations should always reflect specific objectives and questions.

The four categories identified on the basis of methodologies, approaches and scope are:

- Ex-post project evaluation: the focus is primarily on the implementation of the project without isolating the contribution of the intervention;
- In-depth evaluation: the focus is on establishing a causal relationship between the intervention and the impact generated. A counterfactual is used and the analytical scope is broadened to allow for a more in-depth understanding of the change produced;

.....

14 The three bodies that make up French development cooperation are: the Evaluation Unit based in the Ministry of Foreign Affairs, the Evaluation Unit for development activities (UEAD) and the AFD's Evaluation and Capitalisation Divisions; they all report to the ODA and are tasked with conducting or supervising the evaluation of interventions that their organizations manage or implement.

- Strategic evaluations: the framework (geographical, sectoral and cross-cutting) reflects AFD's interpretation of the mission established by supervising authorities;
- Summaries of evaluations: meta-evaluations which "evaluate the evaluation" in which the emphasis is on the evaluation process and its efficiency in relation to knowledge and accountability requirements.

In-depth evaluation in particular, focuses on projects, tools and clusters of projects while using a theory-based method (comparative qualitative analyses, contribution analysis) and making use of experimental and quasi-experimental methods to construct a counterfactual condition (using RCT where possible). It aims to understand the link between results and the actual impact generated by the delivery of the intervention, with a high level of statistical confidence. This type of evaluation largely concerns innovative projects or ones that implement models promoted by the Agency.

In-depth evaluation is based on the belief that the impact of a development project cannot be measured solely by observing the progress made in beneficiary wellbeing due to the fact that there is no way to certify that the change occurred as a result of the intervention or whether it would have happened anyway. Nonetheless, this definition of impact evaluation, heavily focused on the construction of a rigorous counterfactual condition, is not seen to be useful or to be of a sufficiently high quality. For this reason, the Agency promotes a more "in-depth" evaluation aimed at understanding behavior and the contextual mechanisms that link the intervention to the change observed in beneficiary wellbeing. Delving deeper means analyzing the intervention's targets, its context, the ways in which impact was generated and why. This is necessarily followed by a qualitative analysis that is conducted alongside the quantitative component which allows evaluators to go beyond the mere measurement of impact.

Impact evaluations are conducted by the Evaluation Division together with the Agency's operational department which is tasked with selecting projects that are suitable for evaluation. The Division must ensure that conditions exist for the adoption of the most appropriate method and that the results produced are correctly disseminated and employed. Furthermore, there are three levels of quality control that the evaluations of Agency-funded projects must undergo. This applies in particular to NGO projects for which external quality control mechanisms are adopted. In fact, the NGO initiatives that receive funding are also subject to evaluation within clusters of projects that relate to a specific country or sector. These are jointly managed by the "NGO partnership" division and by the Evaluation Division; these evaluations supplement those carried out by the NGOs themselves.



Micro-credit Program in Rural Areas of Morocco (Duflo et al., 2008)

Project area: Access to credit for rural development

Project partners and implementers: a microcredit program led by Al Alamana, one of the largest microfinance institutions in Morocco and financed by the AFD. Between 2006 and 2008 the institute attempted to extend its activities to rural areas, opening over 100 branches.

Project objectives and beneficiaries: to introduce a new microcredit program in dispersed rural areas in Morocco in order to extend the credit limits of households. The program introduced eligibility criteria to exclude families whose income derives solely from agriculture or who have just started up businesses.

Evaluation: This study is part of a “randomized” impact evaluation series that aims to gather findings from microcredit programs. It was jointly carried out by the AFD, Al Alamana e Jameel Poverty Action Lab and assessed the (observable) characteristics of the population that lives in target areas, in order to compare their situation with available household data for similar contexts where there has also been participation in microcredit programs.

Together with the survey conducted across 18 other low-and middle-income countries, this allowed for a comparison between results and the trends registered among comparable populations, and to develop a better understanding of the factors that could improve the program. In fact, the peculiar characteristic of micro-credit programs is the fact that participation is voluntary. As such, it is important to understand the characteristics of potential beneficiaries in order to identify the potential factors that shape the decision to sign up for microcredit. In 2006, surveys¹⁵ were conducted with households from 16 villages (8 treatment groups and 8 control groups) in 7 districts in dispersed rural areas of Morocco for the first round of data collection for the randomized experiment. The data collected on the situation of households before the program was implemented (several weeks earlier) allowed evaluators to observe reactions to the introduction of the program in a context in which formal credit is scarce.

The surveys collected information on demographics, migrations, households, production, consumption, revenues, access to credit and women’s activities. Furthermore, at the beginning of the intervention, weekly records were kept for the households that took part in the program (in treatment villages). These records allowed the effect of time on participation to be measured and for the (observable) household characteristics that encourage program participation to be identified. Since participation is based on self-candidacy, the likelihood of a household joining the program appears to diminish over time. A “proportional hazard” model was adopted in order to measure the effect of time. The assumption at the heart of the impact analysis was that microcredit affects access to credit. Therefore, the causal impact on access to credit is measured simply by comparing the credit transactions of households (which are likely to sign up for microcredit) in treatment and control villages and the credit transactions of households a year after program implementation. The comparison of credit transactions from households that have the highest propensity to become microcredit clients in treatment and control villages reveals clear and significant differences. In fact, access to credit is greater in treatment villages compared to control villages. When microcredit is unavailable (in control villages) household demand is only partially substituted by other forms of credit. This produces the following results:

- In treatment villages, 33% households have obtained credit versus 10% in control villages and
- 30% of families in treatment villages request microcredit compared to 3% in control villages (micro-credit was obtained from other institutions that are not part of the program).

¹⁵ The focus was on “lower-middle class” wealth level in particular, defined on the basis of daily pro-capita consumption that is between \$ 2 and \$ 4 a day. The surveys were carried out twice, once in 2006 and subsequently, two years after the beginning of the program.

Even if demand in treatment villages appears to be relatively limited, households in control villages do not seem to substitute microcredit with other existing sources. This is an important finding for evaluators both in relation to the challenges faced in obtaining credit, as well as the impact of microcredit on extending access to credit.

Findings: The importance that evaluators placed on participation levels in the program is due to the objective of reaching the highest possible number of subjects excluded from accessing formal credit channels, in light of the eligibility criteria imposed by microfinance institutions. After one year, the penetration level of the microfinance institution was 17% of households, in which at least one member was enrolled onto the program; this statistic seemed low albeit in line with the programs assessed (average levels were 18% in rural areas in Developing countries with similar characteristics). The eligibility criteria imposed by the institute excluded households whose income derives solely from agriculture (due to the perceived risks) or start-up businesses. The study highlights the way in which this criteria excluded a large portion of the population in target areas and as such, calls for the need to meet the needs of target population, for example, through individual loan schemes or repayment modalities in line with the production cycle.

2.6. EU International Cooperation and Development

The European Union, together with its Member States, is among the main promoters and funders of sustainable and long-term development across the globe. As a supranational institution with a complex structure and its own approach to development, it cannot easily be compared to the agencies analyzed thus far. The primary actors responsible for evaluation practice are the Directorate-General for International Cooperation and Development – EuropAid – tasked with defining development cooperation policy and the European Union External Action Service (EEAS) which works to align policies and interventions with agreed-upon objectives.

For the European Union, the subject of evaluation in the field of development cooperation is key, both in terms of accountability and for learning purposes, in order to provide greater transparency, improve its own policies and interventions, as well as manage resources without losing sight of long-term impact objectives. It is for this reason that the EU has always been interested and regularly takes part in discussion tables, for example, on the topic of aid (conferences on Aid effectiveness) and evaluation, strengthening its relationship with large international institutions and embracing international evaluation criteria and standards (OCSE).

With regard to the International development cooperation sector, the evaluation policy produced for EuropeAid and EEAS in 2011 continues to be the reference point for partners, beneficiaries (governmental and non-governmental), stakeholders etc.

The document provides a definition of evaluation modeled on that of the OCSE/DAC, namely, of a process that identifies changes as well as positive, negative, intended and unintended effects resul-

ting from the delivery of an intervention. Here, attention is placed on the attribution of change or on the identification of factors that have contributed to generating the change. Evaluation is part of the M&E system which adopts various tools – Evaluation, Results-Oriented Monitoring (ROM) Reviews and Internal Monitoring – to ensure that data collection processes are satisfactory and that performance evaluations meet reporting and learning objectives.

Alongside the definition of evaluation, the criteria and principles of the practice are aligned with those advanced by the OECD. The evaluation principles of transparency, ethics, quality and utility are the most prominent, as well as the “5 criteria” which are further supplemented by *Coherence e Community Value Added*. As a complex structure that is active on several fronts and uses a variety of different channels and tools, the main challenge to evaluation highlighted by the EU concerns the need to uphold clarity and consistency in the case of single interventions, while maintaining the political and operational line shared by Member States. This balancing act is reflected in the evaluation criteria specified for project proposals, which stipulate the need for consistency with macro-objectives and with the interventions of other partners and donors (Directorate General External Relations et al., 2015).

In line with Article 208 of the Treaty of Lisbon, the criteria of consistency allows interventions to be more targeted, structured and effective, as well as avoiding the problem of project overlap and duplication. This, in turn, improves the monitoring of expenditure and increases the efficiency of resource use. The final objective is to produce change and generate impact on living conditions in the most underdeveloped countries, creating added value within communities.

An additional aspect that is highlighted, for example, in the numerous Calls for Proposals published on official websites, concerns the principle of sustainability which includes four aspects: institutional, political, environmental and financial. It is against these principles that the sustainability of impact must be verified, and that the factors which lead to the success or failure of the intervention must be identified.

The Policy presents two different strands of evaluation: one that relates to projects and programs and a strategic strand that relates to long-term evaluation which is carried out on the basis of geographic and thematic criteria.

Evaluation	Scope/Focus	Example of specific focus	User/ Learning Purpose
Strategic evaluations	Corporate issue of strategic importance	Joint programming	Senior managers → to inform strategic choices
	Geographic Thematic / Sector	EU cooperation with a specific Country / Region EU support to a Sector / Theme	Senior managers & managers → to inform strategic choices on programming, and in defining policy
Sector, Project and Programme evaluations	Set of interventions in a specific sector at country level		Operational services Sector / Project / Programme manager
	Individual project / Programme		→ to improve on-going / future sector, project and programme management / design

Source: EC, Evaluation Matters 2011, pag.9

In the first case, evaluation practice is aimed at understanding whether set objectives have been met and identifying potential negative or positive effects in order to reach a full understanding of how the intervention can be improved, allowing the benefits that accrue to an organization and to the institutions and partners involved to be maximized.

For this purpose, the ROM system is recommended as a tool to review the project and provide suggestions as to how the evaluation can capture changes more effectively and subsequently, implement them. In the second case, findings that relate to geographical/thematic policies and the general use of tools are analyzed for a set period of time. Accountability is addressed by measuring the quality of European aid, while identifying recommendations and learning that can inform policy and program design.

The evaluation is managed by the central Evaluation Unit. Alongside these types of evaluations, which include interim, final, and ex-post evaluations, are distinct forms of ex-ante evaluation, namely, appraisals and impact assessments.

Evaluation is split into three different types due to the range of tools that are employed by the aid sector which include, first and foremost, projects, grants and contracts, but also, Budget Support and sector-specific support (European Union, 2014). For example, in the case of projects carried out by NGOs with grant funding, evaluation requirements are stipulated directly in the grant (PRAG,

2016). More generally, the Calls for Proposals published list all the requirements that must be met in final reports, which the funded entity is required to take into account and develop in the course of the project. Since single funded projects come under the umbrella of programs and thematic areas that are set at the European level, there is a tendency to place greater importance on the evaluation of these aspects, rather than on single projects. This also explains why program evaluations and project strategies are published on official EU websites. In this sense, the evaluation of programs is, in effect, a summary of project evaluations or an evaluation of a sample of projects that attempts to establish whether the final objectives of each have been met.



EU International Cooperation and Development, First Report on selected results, July 2013 – June 2014 (EU, 2016).

This document is the first publication commissioned by the EC to improve the accountability of the international development cooperation sector toward European partners (European institutions, civil society organisations, citizens). It presents aggregate results for countries where European development cooperation has funded projects and programs. It aims to provide a snapshot of results, albeit, incomplete, due to the use of primarily quantitative indicators and a large focus on aggregated data. The set of indicators adopted in the evaluation are contained within the EU International Cooperation and Development Result Framework (2015) – an SWD (Staff Working Document) – and their primary aim to strengthen the capacity of the DG for International Cooperation and Development to monitor and report results.

Therefore, the report focuses on projects and programmes that have received more than 750.000 Euros in EU funding and as such, does not include the results of smaller projects. The document is composed of three main sections: 1) Development progress in partner countries, in terms of outcomes and impact; 2) EU contribution to results in partner countries (which includes both aggregated results relating to direct outputs and outcomes, as well as disaggregated results by partner countries and thematic areas); 3) The European Commission's Organisational performance as the manager of EU aid.

In general, aggregated and gender-disaggregated data are used alongside a contribution approach, in order to assess the extent to which the work conducted has led partner countries to meet their development objectives.

The various calls for funding generally focus on the fulfilment of criteria and principles and on the evaluation process, therefore, on the requirements for the Terms of Reference to state the uses of the evaluation, its objective and the composition of its Reference Group. As a whole, the focus that emerges, primarily in the case of projects, relates to the process whereby the project and evaluation are planned. Firstly, in the Inception phase, great importance is given to the logic of the intervention, with a view to constructing the chain of expected results (from inputs to impacts), followed by the questions which the evaluation must answer and the underlying hypotheses. Here, the recommended tool is the New Logic Framework, reviewed in 2015, which assists organizations in their project planning and evaluation activities. In the implementation phase, information can be collected via the use of different instruments, from the analysis of documents, to surveys. Here, careful attention must be paid to the reliability and validity of results which be verified in several ways, including through triangulation.

At the end of the evaluation process, a final report must be submitted, which presents results and evaluation findings, while highlighting the challenges faced and making recommendations for future improvements. Furthermore, in light of the formative objectives of the evaluation which seek to use emerging findings, a participatory approach is favoured and aims to involve stakeholders (defined as NGOs, central or decentralized public bodies) in the different phases of the evaluation process.

With regard to the Budget Support tool, namely, the transfer of financial resources to a partner country in order to promote intensive dialogue and achieve greater impact (European Union, 2004), the evaluation process is managed by the Evaluation Unit Headquarters, due to the greater complexity and strategic nature of the intervention.

When evaluating this type of tool, the OECD document *Evaluating Budget Support* (2012), provides key guidance by presenting a methodological framework and approach to evaluate the use of Budget Support in meeting set outcomes and impact objectives. The methodology proposed includes the Comprehensive Evaluation Framework (CEF) which presents the hypothesized sequence of Budget Support effects across five analytical levels (ranging from inputs to impact) and the Three Step Approach which includes the following steps: 1) the assessment of the the causal links between three elements, inputs, direct outputs and indirect outputs; 2) the assessment of outcomes and impact of Budget Support at the country level and the identification of the main determining factors of outcomes and impact through policy impact evaluation techniques; 3) an exploration of the contribution of Budget Support to national policies, strategies and spending actions.

3. Measurement models for impact evaluation in development cooperation

3.1. Measurement models and classifications

The overview of existing debate on impact evaluation in the context of development cooperation, together with examples of policies and practices promoted internationally, highlights the lack of shared approaches and methodologies to be followed during the evaluation process.

The absence of a common language when measuring the social impact generated by non-for-profit organizations is particularly evident in the field of development cooperation and, to a certain extent, is made worse by the complexity of the specific factors of this field. The focus on Developing and Transition Countries, complex contexts of humanitarian emergencies where organizations are called in to intervene create challenges for obtaining an objective evaluation.

The analysis carried out in previous sections highlighted a series of open questions on evaluation methodologies, but at the same time it allowed to deepen a series of documents adopted by different development cooperation entities in order to illustrate principles, frameworks and recommended and / or implemented models for assessing social impact in international cooperation. The table below summarizes what emerged.

Body	Document Title	Suggested Method
World Bank E Inter-America Development Bank (IDB) (2016)	“Impact Evaluation in Practice”, second edition, Handbook	Quantitative and Mixed methods
USAID (2013)	“Impact Evaluations” Technical Note per staff e partners.	Mixed methods preferable
DFID (2013)	“International Development Evaluation Policy”, Documento di Policy	Mixed methods preferable
InterAction (Bamberger, 2012)	“Impact Evaluation Note” 4 documenti che formano una Guidance Note	Mixed methods preferable
ADB (2006)	Impact Evaluation Methodological and Operational Issues Pubblicazione per ADB	Quantitative or Mixed methods if possible (this combines the benefits of both quantitative and qualitative methods).
Bill & Melinda Gate Foundation	Evaluation Policy (Sito ufficiale)	Quantitative and qualitative
OXFAM (2016)	Effectiveness Review Series, 'enhancing effectiveness throught evidence-based learning' - Manuale	Small n project Community development Project Use of qualitative method and if possible qualita- tive also
WFP (2016)	“Technical Note Impact Evaluation” Technical note	Mixed methods preferable
FAO	Impact Evaluation (Suggerzioni sul sito della FAO)	Mixed methods are preferable (for very small projects, this method is useful even if more diffi- cult in practice; a selection of case studies can be used in this case)

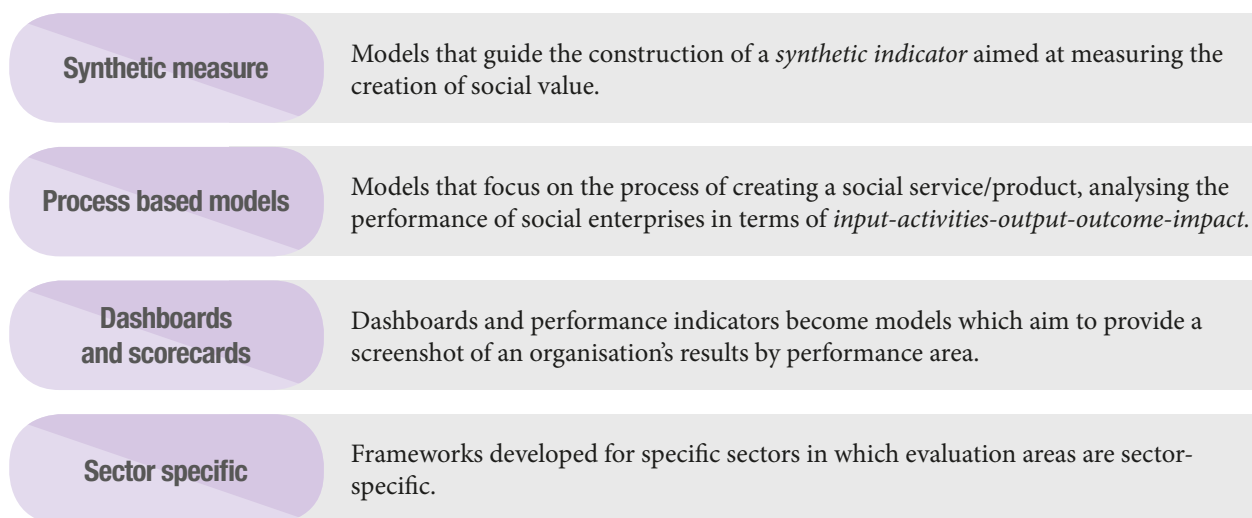
Beyond providing specific guiding principles to be adopted in evaluative practice, the analysis has allowed to focus specifically on the identification of a series of social impact assessment frameworks, particularly suitable for application in development cooperation. A framework of analysis differs from the single methodologies described in previous paragraphs (whether qualitative, quantitative or mixed), as it delineates the phases and steps that make up the evaluation in a broader way. Beyond the indications concerning guiding principles to be adopted, the analysis An analysis framework differs from the single methodologies (be they qualitative, quantitative or mixed) described in the previous paragraphs, as it delineates the phases and steps.

Although a framework may include specific methods, its more general objective is to provide guidelines for conducting the process itself. Frameworks are highly flexible tools that can be easily

adapted to the needs of organizations. In this sense, the perception of evaluation as a very costly process can also be mitigated by the wide range of existing frameworks which vary significantly in their aims, data requirements and analytical perspectives. As a result, organizations can select the most appropriate framework on the basis of their resource availability. The proposal to identify a framework of reference rather than a single tool is consistent with the aim of using evaluation as a learning tool to improve planning from the perspective of *adaptive management*.

Here, evaluation itself becomes a tool for development. In the paragraphs that follow, we review the principal models available, following the classification outlined below: (Bengo, Arena, Azzone, Calderini, 2016):

FIG.1 - Classification of models



It is interesting to note that the most consolidated evaluation frameworks are largely based on the Theory of Change (ToC) model, that describes a sequence of events necessary to achieve a desired change. The ToC supports an organization to consider its planned long-term goal and, from this starting point, to identify the activities that are most conducive to its achievement, as well as the resources required to make this happen. The analytical process is called *backwards mapping* and highlights the causal connections between different elements, allowing the organization to understand the link between the various aspects of its work and its final objective.

The ToC is based on concrete evidence on the effectiveness of a specific intervention and it is strongly influenced by the perspectives of stakeholders whose inclusion in the process guarantees a high degree of validity as causal links are identified on the basis of significant reflection. In this context, the evaluation becomes participatory as stakeholders are involved – in accordance with Pretty's well-known typology of participation (1995) – not only as subjects that are consulted, but also, as interactive actors within the analysis, prior to the planning phase.

3.2. Synthetic measures

3.2.1. Social Return On Investment (SROI)

Social Return on Investment (SROI) is one of the reference models in the field of impact evaluation process.. It was originally developed by the American philanthropic organization, the Roberts Enterprise Development Fund (REDF), with the aim of proposing a methodology that could calculate the non-financial benefits deriving from labor market inclusion activities. SROI is based on the financial concept of return on investment (ROI), a performance indicator that provides information on the number of times an investment is regained by the investor as a result of the activity financed. In the same way, SROI aims to quantify the amount of social impact created, in relation to the cost of investment and expresses social value in monetary terms through a process called “monetization”. It is important to note that value is *monetized* solely as a way to make information on impact comprehensible and comparable through the adoption of a common unit of measurement for investment and social value.

SROI is based on the theory of change and includes the identification of resources, activities, outputs and outcomes necessary in order to create impact. it requires the involvement of all stakeholders in each phase, to establish the primary changes that will affect project/organizational actors, define the results and select suitable indicators in order to verify them. Stakeholders are engaged through a series of qualitative and quantitative methods, such as focus groups, interviews, questionnaires and participant observation.

The SROI consists of six main phases:

1. Establishing scope and identifying stakeholders: defining the activities and objectives that will be analyzed, establishing whether the evaluation will be a provisional one or ex-post analysis; identifying the time frame of the analysis; identifying key stakeholders and how they will be engaged.
2. Mapping changes: identifying the outcomes to be achieved for each stakeholder group that is included in the analysis following stakeholder engagement activities.
3. Giving changes an economic value: converting the qualitative changes desired for stakeholders into quantitative data that can be expressed in monetary terms.
4. Calculating the SROI: in this phase it is necessary to take into account the role of factors that may contribute to create change: deadweight (the amount of outcome that would have happened even if the activity had not taken place), attribution (amount of outcome that was caused by the contribution of other organizations or people), drop-off (the deterioration of impact over time).

5. Calculating the SROI ratio: this is derived from the comparison between the value of outcomes which are monetized using financial proxies and the total value of the investment made. The result is a ratio such as 3€:1€ which means that for every 1€ investment in the project 3€ of social impact is generated.
6. Dissemination of results: both internally and externally.



La Bussola Daycare center, Padua

La Bussola is a daycare center in the city of Padua which, since 2005, has been providing initial reception services to individuals facing severe social exclusion and in particular, homelessness. The services are managed by the Gruppo R cooperative, in partnership with the *Rogationist Congregation*, the Municipality of Padua the Paduan Caritas pastoral organization. In January 2014, the “Ritorno al Futuro” project was launched with the aim of supporting individuals who access tailored support for social, employment and personal development services. The project was evaluated by Human Foundation using the SROI methodology which led to the identification and analysis of the impact created by the project and to an estimation of the financial value of the social change generated. In particular, the evaluation found that an investment of 117.500,00 € generated a social value of 753.282,41 € indicating that every euro invested in the project would deliver 6.41€ over a five-year period.

For further details please see: *Human Foundation*, 2015.

3.2.2. Local Multiplier 3 (LM3)

LM3 was developed by Sacks in 2002 for NEF and is based on the concept that an economic multiplier effect measuring the impact of expenditure on the economy, taking into account knock-on effects. The *multiplier effect* is an economic concept which states that the higher the proportion of money spent in the local economy is, the larger the income generated for the local population. More income locally or nationally means more jobs, higher pay and more tax revenue for government, resulting in improved living conditions for people locally. The measurement process involves three phases:

- Identifying a revenue source
- Monitoring the way revenue is spent
- Monitoring the way revenue is re-spent within a specific geographical area (local economy).

3.2.3. Gamma Model

In 2011, Grabenwarter and Liechtenstein proposed the Gamma Model which originates in the field of impact investing and provides a synthetic indicator able of measuring social and environmental impact in line with financial market standards. The model is based on the *Capital Asset Pricing Model* (CAPM) and the gamma factor integrates the concept of social and/or environmental impact into the overall return of an investment. As such, the Gamma Model is a synthetic measure of the value created by an impact investment. In particular, the model divides impact assessment into impact objectives at the investment level which become performance indicators (KPIs) and returns of an investment which become “impact-adjusted return”. Furthermore, the model proposes the use of impact indicators as a starting point for the performance indicators of impact investments. These indicators measure the quality of investment decisions, based on impact findings. This provides an integrated measure of financial performance and impact performance at the portfolio level, allowing single organizations to select significant KPIs.

3.3. Process Based Models

3.3.1. Process Tracing

Process Tracing (PT) is a qualitative approach to social impact evaluation developed in response to the promotion of the quantitative and counterfactual method in the context of international development cooperation. PT fits into a framework aimed at researching models that, although rigorous, can also be applied successfully in cases where time and resources are limited and/or there is a lack of detailed information and reference data that can support counterfactual evidence. The method allows to measure the ex-post changes that have occurred, identifying a series of causal links without relying on control groups. The principle of this tool is the identification of the main steps - that begin with a hypothetical cause (e.g. an intervention for territorial development) – by leading to a precise outcome. Similarly to the Theory of Change, PT requires the overarching causal mechanism that connects these two elements to be “unpacked”, in a process that is similar to the one carried out by an investigator, in which evidence is gathered to demonstrate the strength of the links hypothesized, through probability tests.

PT involves five phases (Beach and Pedersen, 2013):

1. Developing a hypothetical causal mechanism: starting from the project's theory of change, identifying the details and defining all the steps that begin with the intervention and lead to the desired outcome;

2. Making the causal mechanism operational: identifying the empirical explanations and observable data that confirm each of the steps hypothesized;
3. Collecting the data: primary and secondary data to obtain observable proof of each step;
4. Evaluating the strength of evidence: applying one of four tests developed for this phase ('straw-in-the-wind' test, 'hoop' test, 'smoking gun' test, 'doubly decisive' test¹⁶).
5. Establishing process conclusions: On the basis of the robustness of each hypothesis, establishing whether the overall causal model should be accepted or rejected.



Universal Health Campaign in Ghana

Process Tracing was used in the evaluation conducted by the Universal Health Campaign in Ghana, an advocacy campaign led by civil society organizations to advance free and universal access to health care, funded by Oxfam GB until 2013. The evaluation process commissioned by Oxfam GB and conducted in 2012-2013 (Stedman-Bryce, 2013) was based on a PT protocol developed by Oxfam GB (2011). The protocol included the verification of the links between the campaign and a series of outcomes on the basis of three elements: a) short-listing one or more explanations for the outcome, b) ruling out alternative competing explanations that were incompatible with the empirical explanations put forward, and finally, c) estimating the level of influence that each of the explanations put forward had in creating the change in question.

For further details please see: *Punton and Welle, 2015*.

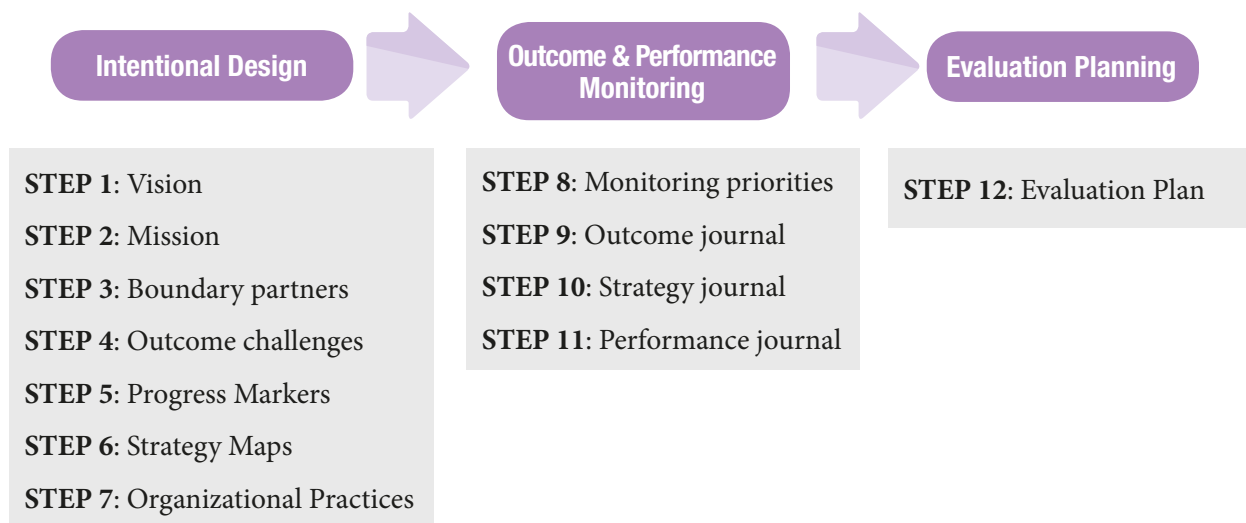
3.3.2. Outcome Mapping

Outcome Mapping (OM) is a framework developed by the International Development Research Centre (IDRC) and a methodology that devises evaluation plans, which allow organisations to record their results and the change they have generated. The framework emerged in response to the confusion surrounding the very concept of “impact” which is often misunderstood by organisations in the context of the evaluation process. In response to this barrier, OM supports organizations to shift their focus from project outputs to outcomes, defined as tangible changes in behavior, relationships and in the activities or actions of all the individuals directly involved in a specific

programme (Earl et al., 2001). The specific focus of an evaluation that uses the OM methodology is not on the change itself as much as on the contribution made to creating it, taking into account the effective role of the organization in question, which works together with different actors toward the same aim.

Within the OM methodology, an evaluation plan is defined through 12 steps and 3 macro stages:

1. **Intentional Design:** establishing the change that a project aims to bring about and planning the specific activities to be carried in order to maximize the chances of success.
2. **Outcome & Performance Monitoring:** identifying monitoring and evaluation priorities as well as defining a system to verify the progress made towards planned results and outcomes. Progress is tracked through a series of indicators defined in Phase 1 (STEP 5: Progress Markers).
3. **Evaluation Planning:** on the basis of the elements defined in the first two phases, the evaluation plan is developed and presented.





In Madagascar, the Swiss development cooperation agency together with its local partners decided to improve the SAHA program for rural development, by shifting the focus away from the micro-level, towards the me-so-level. In this sense, it moved away from working directly with small agricultural organizations and began to develop partnerships with intermediary organizations that were responsible for the planning and implementation of activities. The evaluation of this project was conducted using the OM methodology and focused not only on the changes that occurred in conditions (for example, improved infrastructure and services), but also on those regarding the mindset and behavior of workers. It shed light on the fact that the decentralized nature of responsibilities among the local population and institutions is one of the primary critical success factors of the project.

For further details please see: *Hearn, 2016*.

3.3.3. Methodology For Impact Analysis And Assessment (MIAA)

The Methodology for Impact Analysis and Assessment (MIAA) is a methodology outlined by Investing for Good in the book *The Good Analyst* (Hornsby, 2012). The framework includes three evaluation areas: Mission Statement, Beneficiary Perspective and Wider Impact. Each of these is composed of a series of sections containing specific indicators, which an organization can use to evaluate its actions and the resulting impact on the community via a three-level scale: Low – Medium – High. MIAA also includes a final scorecard which assigns each indicator a numerical score. In this way, the score of an organization can be compared to the maximum score obtainable in each area and on this basis, the areas of strengths and weaknesses can be evaluated.

FIG.2 - Example of a scorecard to calculate overall score awarded

	IMPACT	POTENTIAL SCORE	SCORE AWARDED
1	MISSION FULFILMENT	40	
1.1	Mission Statement	2	
1.1.1	Mission Statement	2	
1.2	Context and Focus	3	
1.2.1	Understanding the Problem	2	
1.2.2	Understanding Beneficiaries	1	
1.3	Impact Activities	15	
1.3.1	Theory of Change	3	
1.3.2-a	Impact Measurement: Use of Appropriate Indicators	2	
1.3.2-b	Impact Measurement: Quality of Data	1	
1.3.2-c	Impact Measurement: Target and Objectives	1	
1.3.3-a	Impact Reporting: Transparency	2	
1.3.3-b	Impact Reporting: External Validation	1	
1.3.4-a	Balance and Alignment: Congruence	3	
1.3.4-b	Balance and Alignment: Attitude to Profit	1	
1.3.4-c	Balance and Alignment: Mission Drift	1	
1.4	Results	12	
1.4.1-a	Results: Delivery of Impact	5	
1.4.1-b	Results: Targets and Objectives	2	
1.4.1-c	Results: Performance Improvement	2	
1.4.2-a	Accreditation and Comparison: Accreditation	1	
1.4.2-b	Accreditation and Comparison: Class Comparison	2	

For further details please see: *Hornsby, 2012*

3.3.4. Social Impact Assessment (SIA)

In 2012, the Global Social Venture Competition (GSVC) developed the Social Impact Assessment (SIA) tool. SIA uses the logic put forward in the SROI guidelines to define, measure and record the impact created but restricts the analysis to three primary results that the organization aims to achieve, without tying the analysis to the monetization of outcomes.

In fact, social impact evaluation involves the following three steps:

- Defining the social value created by an organization through the theory of change;
- Identifying, through stakeholder engagement activities, the three primary indicators of social impact that are most linked to the social results desired and defining the methodology of calculation;
- Planning the impact measurement process and the continued monitoring of the unintended and/or negative effects of the activities, in a clear and feasible way.

In order to guide the selection of indicators, the model makes use of the *Impact Reporting Standard & Investment (IRIS)* framework which provides reporting standards for social and environmental performance. More specifically, IRIS provides a dashboard of indicators with defined standards that that can be applied across sectors (GIIN 2011). In this way, the taxonomy allows an organization to:

- Preempt the social changes produced by a project or a policy;
- Measure the impact and systematize the data;
- Assess the potential social implications;
- Communicate the changes to the community;
- Define the steps required in order to mitigate negative effects and enhance positive ones;
- Identify the benefits that derive from changes produced which do not improve the condition of certain stakeholder groups.

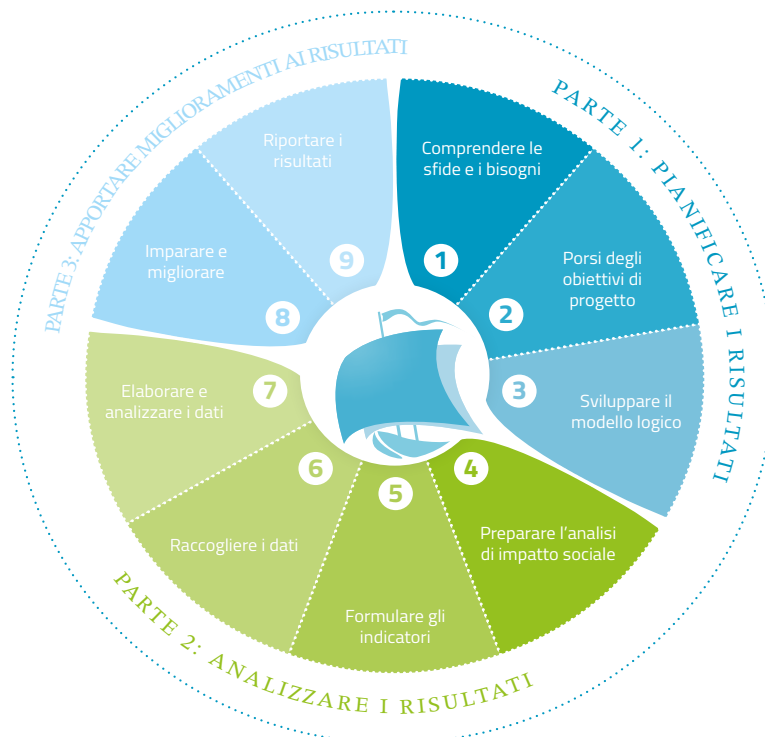
3.3.5. Impact Navigator

The Social Impact Navigator is a guide produced by the think tank PHINEO with the support of the World Bank's Global Partnership for Social Accountability (GPSA) and Bertelsmann Stiftung.

It begins by presenting the expertise of over 800 civil society organizations and offers tools, advice and practical examples for non-for-profit organizations to measure their results in a more effective way, as well as to obtain improved results and share their work and impact with people all around the world. The structure of the guide is based on the description of the management cycle of impact-oriented projects and involves three stages:

- Planning results;
- Analysing results;
- Improving results.

The guide provides practical steps to integrate impact into the project planning phase and to approach the question of evaluation both within the organization and externally.



Source: *Social Impact Navigator*, Phineo

3.3.6. Developmental Evaluation

Developmental Evaluation (whose principles are similar to those of Action Evaluation, Adaptive Evaluation, Emergent Evaluation and Real time Evaluation approaches) is an approach rather than a single evaluation method aimed at supporting the development of a program, product, staff and/or organizational structure. Developed in recent years, it provides feedback and produces knowledge to support or readapt the process of change in real-time, adjusting strategy and activity planning while delivery is underway. It stands out from other methods due to the fact that the approach includes the use of new measure tools and monitoring tools that are developed on the basis of emerging and evolving objectives. The proposed methodology is recommended in situations where traditional evaluation models are not easily applied, namely, in development contexts, in cases where there is a high degree of innovation, of uncertainty and dynamism, where systemic changes are being produced and where adaptive management is required. In particular, it is an evaluation approach suited to the following three complex conditions: difficulty predicting and planning; highly dynamic and rapid change; the existence of interdependent and nonlinear relations that make it difficult to trace cause-effect chains and where the result is highly dependent on the context.

According to Developmental Evaluation, learning and capacity-building *of organizations* should be the specific focus of the evaluation process. This is why evaluation feedback must be given within a set time frame (for this reason, also named real-time evaluation) and not after a long and conclusive evaluation process, to enable resource to be redirected toward outcomes that have emerged during project delivery, if necessary. The evaluator must be flexible, creative and adaptable, and become a veritable project team member, whose specific role is to facilitate rigorous reflection based on the evidence that emerges from the evaluation, informing the decisions made regarding the project and its development objectives. Rather than being an outward-looking principle, accountability is about project actors and stakeholders who promote and are involved in generating change respecting the fundamental values of commitment and learning. The aim of Developmental Evaluation is to provide data that will lead to a strategic response, rather than promoting externally-driven inspection through which the potential gaps and failures of an initiative are identified. There are several evaluation methods that can be adopted as part of this approach, all of which may be fitting and in line with the characteristics and dynamics of a specific situation (Patton, 2011): by their very nature, qualitative methods are more appropriate tools, due to their greater flexibility and ability to conduct an in-depth analysis of the relations and processes underway.

3.4. Dashboards and scorecards

3.4.1. Fit For Purpose

Fit For Purpose (FFP) is a *self-assessment* tool developed by the Development Trust Association for *community enterprises*, however, it is particularly suited to all non-for-profit organisations whose objective is to increase the socio-economic development of the community. FFP is similar to a diagnostic tool and provides organizations with an analytical grid to evaluate the work conducted in relation to set objectives and allows areas of success and improvement to be identified. Although it began as a tool for self-evaluation, it can also be used to guide external evaluation.

FFP is composed of 5 sections: governance, business planning, financial management, partnerships, policies and procedures. A series of result indicators are defined for each of these sections in which the organization (or external evaluator) assesses progress on the basis of three criteria: reached, partially reached or not reached. Organizations should use the model to uncover objectives that have not yet been met and require further effort. Additionally, every indicator must be accompanied by documentation to support the evaluation results.

FIG.3 - Example of indicators for Section 1, Governance

● Met
 ● Partly met
 ○ Unmet

INDICATOR	GUIDANCE	RATING/NOTES
1.1 The make-up of the board reflects the mix of the community it serves; the board is accountable to that community.	How does the organisation ensure that the make up of the board reflects the mix of the community? How does the structure of your organisation ensure accountability to the community?	● ● ○
1.2 The organisation has a clear and agreed mission, aims and definition of its target beneficiaries which have been defined through community consultation/mapping exercise/evidence of need.	What is your organisation seeking to achieve? Who benefits from this? How is this understood by all concerned?	● ● ○
1.3 The organisation's governing documents allow it to carry out its trading activities.	Is the power to trade stated in the organisation's "constitution"? If the organisation carries out primary purpose trading, does it meet Charity Commission requirements?	● ● ○
1.4 The organisation's governing documents prevent personal profit being made by members of the Board.	Does the organisation's "constitution" prevent Board members profiting from their involvement in the organisation?	● ● ○
1.5 The organisation can demonstrate that any surpluses generated by the organisation and any trading subsidiary are used for community/public benefit or re-invested back into the organisation.	Can you point to a clause within your governing documents which set this out?	● ● ○
1.6 The Board makes all policy decisions.	How does your organisation ensure that the Board ultimately controls policy?	● ● ○
1.7 The organisation keeps the community informed of its activities and welcomes and acts upon input from the community	How does your organisation ensure the flow of information to and from the community?	● ● ○

For further details please see: *Development Trust Association, 2008*

3.4.2. Charity Analysis Framework

The Charity Analysis Framework (CAF) is a grading grid to evaluate the work of organizations. It was developed by New Philanthropy Capital (NPC) and follows the same structure as the Fit For Purpose tool. It is a grid that organizations can administer themselves or use for an external evaluation in order to evaluate their work in connection with set objectives and highlight strengths and weaknesses. The CAF is composed of six areas of analysis: activities, results, leadership, people & resources, finances and ambition. For each of these areas, charities receive one of four ratings: excellent, good, satisfactory, below expectations. The model was developed for the evaluation of charities however, it can also be adapted to different contexts, in particular, to: large organizations, awareness-raising activities, community organizations, associations and organizations that work in Developing Countries.

FIG.4 - Examples of Evaluation criteria

AREA FOR ANALYSIS	ASSESSMENT CRITERIA (ABSTRACT)	GRADE
Activities	<ul style="list-style-type: none"> • Focus on greatest needs • Ability to adapt and innovate • Links between activities 	
Results	<ul style="list-style-type: none"> • Evidence of positive results • Quality of evidence • Sharing results 	
Leadership	<ul style="list-style-type: none"> • Management Team • Governance • Vision and Strategy • Priorities 	
People and Resources	<ul style="list-style-type: none"> • Staff • Use of volunteers 	
Finances	<ul style="list-style-type: none"> • Financial security • Efficiency • Unit costs 	
Ambition	<ul style="list-style-type: none"> • Potential to grow • Potential to replicate • Potential to improve results 	
		Excellent
		Good
		Satisfactory
		Below expectations

For further details please see: *Copps e Vernon, 2010*.

3.5. Sectoral Frameworks

3.5.1. Poverty And Social Impact Assessment- PSIA

Poverty and Social Impact Assessment (PSIA) is a framework developed by the World Bank to measure the social impact of policy reforms and trace its geographical distribution, as well as to increase the wellbeing of vulnerable groups in Developing Countries. In fact, in order to systematically assess the social implications that these reforms have on poverty reduction in target contexts, specific tools must be adopted. To this end, the framework provides a series of key principles, tool and methods that can be adopted by practitioners in order to evaluate the impact created (The World Bank, 2003). The PSIA is based on a series of key questions that guide the evaluation process, specifically:

- (what) What is being analyzed?
- (what) What is the welfare measure being assessed?
- (who) Whose welfare is being analyzed?
- (how) How are impacts channeled?
- (how) How do institutions affect outcomes?
- (when) When do impacts materialize?
- (if) What are the risks of an unexpected outcome?

On the basis of these questions, the PSIA proposed the following 10 elements that make for good evaluation practice:

1. Asking the right questions by clearly identifying the scope of the analysis (understanding the questions being posed, identifying the policy reform to be analyzed, formulating research hypotheses, and identifying policy constraints);
2. Identifying relevant stakeholders, namely, the people, groups of organizations that are affected by the policy and those who can affect policy implementation;
3. Understanding impact transmission channels, identifying basic hypotheses and links;
4. Understanding the role of institutions, focusing particularly on the structure of the market and on development agencies that operate locally;

5. Gathering data and information;
6. Analyzing impacts, selecting the most appropriate model (several are proposed) on the basis of time and data constraints, as well as on available resources and skills;
7. Contemplating enhancement and compensation measures;
8. Assessing risks;
9. Monitoring and Evaluating impacts generated both during implementation and in the ex-post phase;
10. Fostering policy debate and feeding back into future policy choices.



The reform to privatize cotton in Chad

The government of Chad used PSIA to guide the reform to privatize Cotonchad, the parastatal organization that manages national cotton production. The objective of the reform was to overcome a series of structural inefficiencies that had led to progressively low yields of cotton – a crop that has traditionally been key in Chad due to the revenue it generates as well as its role in poverty reduction. The ex-ante analysis conducted was a preliminary context analysis that aimed to identify the problem and to define the assumptions to guide the reform. The evaluation team worked closely with key stakeholders and identified the areas in which the reform was likely to improve the performance of the sector. In line with these findings, the inputs required to achieve planned results and outcomes were established, allowing the final objective of the reform to be met. The team specifically identified the causal links between the elements for each transmission channel of the reform.

For further details please see: *The World Bank*, 2003.

3.5.2. Health Impact Assessment - HIA

The Health Impact Assessment (HIA) is a framework developed by the World Health Organization (WHO) which proposes a series of procedures, methods and tools to estimate the potential effects of a project on the health of a population and in particular, on disadvantaged and vulnerable groups. The framework is largely prospective and aims to produce a series of recommendations for decision-makers and stakeholders in order to maximize the positive impacts of a proposal and mitigate negative impacts. The framework provides a series of qualitative, quantitative and participatory techniques through which to estimate impact and involves five phases:

1. Screening: determining whether the conditions exist for HIA to be applied (analysis of the project/program/policy and of its potential health effects);

2. Scoping: identifying the scope of the analysis (who to involve, how and when to carry out the evaluation, which resources to use, what the primary object of the analysis is);
3. Appraisal: undertaking the evaluation while identifying the potential risks for health and gathering proof of their impacts;
4. Reporting: producing a series of recommendations in order to reduce risks tied to the project and safeguard as well as improve the health of the population.
5. Monitoring: assessing whether the HIA has effectively influenced decision-making and whether the recommendations have been taken into account.



Morice Town Home Zone

Home Zones is an initiative promoted by the British Institute of Highway Incorporated Engineers aimed at re-generating residential areas and transforming them into environments that are safe for all citizens. The Morice Town project (Plymouth, United Kingdom) constitutes the first area to be assessed using the HIA framework among the Home Zones initiatives. In this sense the local community was successfully included in the entire evaluation process. The specific object of the evaluation was: the quality of life, pollution, entertainment, security, transport and health of population.

For further details please see: *Elliston e Maconachie*, 2002.

The figure below contains a brief summary of the frameworks presented thus far:

FIG.5 - Models/Actors Combination



4. Methodological proposal for evaluation

4.1. Choice of evaluation model/framework

While the previous chapter presented existing evaluation models/frameworks, the following chapter aims to support organizations in the choice of models/frameworks to adopt, offering them step-by-step guidance on implementation in line with Project Cycle Management (PCM), the primary project management tool in the field of development cooperation. This section of the report will therefore integrate impact evaluation with the activities and modes of project management that organizations have already put in place, reviewing the elements of the project cycle (process) from its inception through the lens of impact measurement. The idea that organizations can choose particular models and potential combinations of measurement and analytical methods, stems from the assumption that there is no single way to conduct evaluation, just as there is no single perspective from which to ‘observe’ the world. For this reason, the following chapter does not provide a standardized evaluation model, as the “one size fits all” philosophy is deemed to be inappropriate in this context. Conversely, within the field of international development cooperation, the standardization of the evaluation process seems more appropriate, starting with the choice of which methodology to use. If the evaluation process put in place is comprehensive and robust, it will certainly include all the elements appropriate to the development cooperation sector. (E.g. sustainability, empowerment, growth in beneficiary population). Thus, the choice of which methodology to adopt is often tied to the beneficiary type (applicant) of the evaluation and will meet its set objectives, reflecting the values and principles espoused by the actors involved. For example, the use of methods that include participatory techniques is coherent with a specific concept of development. Furthermore, an additional factor that influences choice is tied to the scope of the analysis, which can confine the evaluation to a specific project or broaden it out to include an entire organization. This report will focus on the project level.

An additional methodological note concerns the opportunity to measure social impacts, which, by definition, can be captured and therefore correctly measured within a long-term perspective. For a long-term perspective to be adopted, the evaluation must exceed the implementation time frame and therefore, the end of the development project. In order to conduct the evaluation, a monitoring system should be put in place to cover a period that spans between 3 and 5 years *after* the project has ended. This approach cannot be implemented exclusively by the organization delivering the project but rather, must become a shared practice that is led and supported by funding agencies

that allocate resources specifically for these purposes.

To date, the system as a whole does not support organizations to “go back to” their results over time; for this reason our first objective has been to capture and therefore measure outcomes, which are currently the best proxies for the potential impact that can be generated in the future.

Given that the objective of project evaluation is *measurement*, the organization delivering the project has two aims: firstly, internal growth, in a broad sense, for both the organization and all the subjects involved in the initiative being evaluated and secondly, to improve its external credibility.

A growing number of non-for-profit organizations are beginning to analyze and record the effects generated by projects. Nonetheless, this practice is not entirely consolidated and it is generally seen as a significant challenge. Often, interest in the topic of impact is largely guided by the need to present results externally or to obtain legitimacy. In any case, the **essence of an approach that embraces impact is its interest in learning and in the continued improvement of the work being carried out**. Only organizations that possess in-depth knowledge of the results they produce and of their areas of strengths and weaknesses can use this knowledge for improvement and to move closer to achieving their objectives.

In fact, the evaluation process allows an organization to increase its internal competencies, through the:

- Improvement of strategic planning to support decision-making;
- Improvement of processes and, as a result, the strengthened social impact for beneficiaries;
- Maximization of the efficiency and effectiveness of organizational activities/projects;
- Strengthening of reporting capacity;
- Increased engagement and involvement of stakeholders;
- Strengthening of growth and learning across stakeholders groups involved;
- Strengthening of team motivation thanks to the identification of aims that can be objectively verified;

With regard to the specific audiences of evaluation, these can be internal actors, namely, within an organization (*decision-makers, project managers, project workers*) who carry out the activities and the evaluation itself. Their objective is to find indicators and measurement tools to support internal decision-making, improve the efficiency and efficacy of activities and consequently, social results. The application of participatory methods presupposes a joint role with beneficiary target

groups and with stakeholders that represent the community of reference in the project planning and implementation phases. These actors can be considered internal to the delivery of the project and as such, together with project staff and management, have a direct interest in the evaluation. The learning that follows an evaluation process ensures that the consistency and intensity of impact is strengthened, namely, that the ability to respond to the needs of the target population is increasingly effective and that the highest number of direct and indirect beneficiaries are reached and engaged. These are both priorities for interventions across the field of international development cooperation.

With regard to external audiences, alongside the increased credibility of an organization and the improved ability to communicate results attained, evaluation leads to greater transparency and accountability and as a result, attracts capital and investment by demonstrating the effective social return on investment. Notwithstanding the necessary distinctions between private and institutional donors, as well as private and institutional investors, funders have similar evaluation needs. In the case of donors, the objective is to gain knowledge of the social impact generated by an organization/project and the cost of achieving it, while in the case of investors, knowledge is required of investments in socio-economic terms so that decisions can be made on where and how much to invest. For both groups, to date, the separate or aggregated use of synthetic and/or processes-based models appears to be the most consistent and widely adopted solution.

With regard to policy-makers, whether they are lawmakers or public funders that promote contracts and partnerships, the first group requires impact measurement in order to evaluate social policies while the second needs information on the quality of the services provided and the impact generated by the socially-oriented projects that receive funding.

Once the audience and accompanying evaluation objectives and information needs have been established, the appropriate measurement method can be selected on the basis of the characteristics of model clusters analyzed in the previous chapter. This is summarized in the figure below:

FIG.6 - Actors and corresponding models

	STAKEHOLDERS	BISOGNI	MODELLI
FUNDERS	Private or institutional investors	To know the impact of investments in socio-economic terms in order to make decisions on where and how much to invest.	Synthetic measure
	Donors	To know the social impact generated by an organization and the costs of achieving it.	Process based models
POLICY MAKERS	Lawmakers	To know the impact of social policies	Synthetic measure Process based models
	Funders through contracts or partnerships	To know the quality of the services provided and the impact generated by socially-oriented projects.	Process based models
DECISION MAKERS	Subjects that are internal to an organization (project managers, project workers)	To know the indicators and measurement tools required to support internal decision-making, improve the efficiency and effectiveness of activities and improve social results.	Process based models

Additional guidelines for choosing an evaluation model are based on the characteristics of the evaluation itself. Approaches that are too complex and onerous, albeit sophisticated, do not make for a feasible, sustainable and replicable evaluation. Furthermore, the model should account for the range of objectives (social, economic, environmental) pursued by an organization which, as is well-known, produces value for different stakeholders. The model must allow for the constant realignment to the mission of an organization, minimizing the risk of “mission drift”. In addition, it must allow specific project indicators to be easily identified and populated, as well as being consistent with the monitoring system put in place in order to preserve the intervention’s multi-subjective nature. Finally, the design of sector-specific models may prove useful when choosing more appropriate indicators to integrate into non-sectoral models. In this sense, the approach to impact evaluation in development cooperation projects should possess the following characteristics:

- Support the learning and development of the organization;
- Support the empowerment and development of the target community;
- Value the contribution of all actors involved (beneficiaries, partners, institutions and communities);

- Ensure that a positive impact is generated for beneficiaries and that it is used to frame their needs;
- Adopt a method of analysis that is proportional, sustainable and replicable;
- Adopt mixed method quali-quantitative research designs;
- Account for the multiple objectives of an intervention;
- Measure the impact on different stakeholders;
- Account for the alignment between the social mission and results;
- Identify impact indicators that are specific and relevant;
- Carry out the evaluation in line with the monitoring system;
- Communicate impact to investors and donors;
- Communicate impact to beneficiaries;
- Communicate impact to the community and to policy-makers;
- Account for the nature of the intervention (development, reconstruction...).

4.2. Review of impact evaluation models in relation to international development cooperation project needs

The following section evaluates the models presented in Chapter 3 in relation to the criteria outlined in previous paragraphs. The objective is to guide organizations in their choice of appropriate methodologies on the basis of priorities and capacity. The table below was created by the authors to present the different models while indicating their degree of correspondence with each criterion, where “high” means that the model meets the criteria well, “medium” means that it partially meets the criteria, and “low” means that it does not adequately meet the criteria. The assessment carried out suggests that the criteria are best met by “process-based” models, such as Outcome Mapping, SIA, the Impact Navigator and the sectoral PSIA model. Therefore, in the paragraphs to follow, SIA and the Impact Navigator are proposed as guides to the development of evaluation activities. It is important to note that the use and integration of multiple models is often the best choice, for example, synthetic indicators or LM3 on their own would not be adequately evaluate the organization and its activities, however, when used to synthetically evaluate the transition

from outcomes to impact, they may allow for the quantification of impact in monetary terms. We do not believe that everything should be quantified and/or monetized, however, we feel that in the choice over whether to define impact as a long-term change that affects the community, the use of synthetic indicators should be accompanied by the use of qualitative indicators. The adoption of monetization as a technique for quantification, regardless of whether the chosen approach is based on the degree of attribution for outcomes or on impact understood as long-term effects, must be tied to the specific mission of a project and use proxies that are recognized at the country level. Eurofund's quantification of the cost of a NEET individual to the Italian state is an example of a recognized proxy. A similar process is easily carried out, for example, with regard to the quantification of the savings to public spending generated by prevention projects or the savings and increased revenue that follow entry into work programs, which commonly use the average cost of being unemployed. Within other sectors, the process is more complicated, particularly in the field development cooperation in which impact is generated across several areas. Here, this method is more difficult to apply and risks under-estimating the real value generated by organizations.

FIG.7 - Elaborazione propria

	1. SROI	2. LIM3	3. GAMMA	4. PROCESS TRACING	5. OUTCOME MAPPING	6. MIAA	7. SIA	8. IMPACT NAVIGATOR	9. DEVELOPMENTAL EVALUATION	10. FIT FOR PURPOSE	11. CAF	12. PSIA	13. HIA
Support the learning and development of the organization	HIGH	LOW	MEDIUM	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	MEDIUM
Support the empowerment and development of the target community	HIGH	LOW	LOW	LOW	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM
Value the contribution of all actors involved (beneficiaries, partners, institutions, communities)	HIGH	LOW	MEDIUM	MEDIUM	HIGH	MEDIUM	MEDIUM	HIGH	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM
Ensure that a positive impact is generated for beneficiaries and that it is used to frame their needs	HIGH	LOW	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	HIGH	HIGH

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
Adopt a method of analysis that is proportional, sustainable and replicable	MEDIUM	LOW	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	LOW	LOW	MEDIUM	MEDIUM
Adopt mixed method qualitative-quantitative research designs	MEDIUM	LOW	LOW	MEDIUM	MEDIUM	LOW	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH	HIGH
Account for the multiple objectives of an intervention	MEDIUM	LOW	MEDIO	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	MEDIUM
Measure the impact on different stakeholders	HIGH	LOW	MEDIUM	MEDIUM	MEDIUM	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH	LOW
Account for the alignment between the social mission and results	HIGH	LOW	MEDIUM	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH	MEDIUM
Identify impact indicators that are specific and relevant	MEDIUM	LOW	LOW	LOW	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH	LOW
Carry out the evaluation in line with the monitoring system	MEDIUM	LOW	MEDIUM	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH	MEDIUM
Communicate impact to investors and donors	HIGH	HIGH	HIGH	LOW	MEDIUM	MEDIUM	MEDIUM	MEDIUM	MEDIUM	LOW	LOW	MEDIUM	LOW
Communicate impact to beneficiaries	MEDIUM	MEDIUM	MEDIUM	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	HIGH	LOW	LOW	HIGH	MEDIUM
Communicate impact to the community and to policy-makers	MEDIUM	MEDIUM	MEDIUM	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH	HIGH
Account for the nature of the intervention	LOW	LOW	MEDIUM	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	MEDIUM	MEDIUM

4.3. The approach to evaluation and its integration with Project Cycle Management (PCM)

Once the scope of the evaluation (project), the needs of the different actors involved (internal decision-makers, policy-makers and funders) and the methodology most suited to organizational needs and project characteristics have all been established, an evaluation design that is aligned with the aims of development cooperation can be constructed. As a starting point, it is necessary to take into account the approach that is most widely used in project planning within international development cooperation interventions: **Project Cycle Management (PCM)**. PCM is a tool introduced by the European Commission at the beginning of the 1990s to improve the quality of project planning and management. PCM includes a phased approach to the planning and implementation of projects and for each of its phases, assigns roles to the different actors involved, specifying the decisions to be made. In 2015, the Commission reviewed the tool and integrated the logic of medium-to long-term impact evaluation into PCM. This is evident in its inclusion of the **Theory of Change (ToC) and the social value chain**, alongside a more traditional method tied to **Project Cycle Management** used to define and manage projects, which is therefore more operational in nature and connected to medium-to short-term evaluation projects.

Furthermore, past assessments have found approaches based on the value chain to be the most commonly employed and easily aligned to the logic of projects. Reflections that emerged subsequently and deserve attention, concern the potential to create hybrid models based on the value chain (SIA; Impact Navigator...) using synthetic indicators (SROI, LM3) and to integrate the latter with counterfactual analysis. Here, organizations will require additional support in order to make evaluation effective in regard to project planning and impact evaluation.

To this end, the characteristics of both approaches (PCM, ToC – social value chain) are presented below, alongside their commonalities and the steps required to integrate and utilize both approaches. The Theory of Change is more focused on the changes that occur in the lives of the individuals involved or those that occur within a system (outcomes), while PCM places greater importance (if the “General objective” is disregarded) on evaluations that are specifically tied to outputs, on the results of set objectives and on the actions required to meet them.

The Dutch embassy requested that a Theory of Change be constructed for a project that aimed to strengthen the resilience of vulnerable populations in Burundi, in particular, children under the age of 2 and pregnant women. On this basis, an evaluation model was built involving the following steps:

- Impact was identified as the reduction of malnutrition in children under 5 through the delivery of specific nutritional interventions, with a particular focus on the first 1,000 day window;
- 4 medium-term results that lead to impact were identified, including the access to community intake services;
- 3 short-term results were identified as being necessary to the achievement of medium-term results such as: community engagement around the proposed nutritional methodology (level 1), training of community actors for intake (level 2) and the activation of a screening system and of nutritional centers (level 3);
- Activities were established for each short-term result.

In addition, the prerequisites and primary barriers and challenges to address were identified.

The main innovation brought by a PCM approach that seeks integration with the Theory of Change, is the new version of the LOGFRAME MATRIX, the principal Logical Framework Approach (LFA) tool. More specifically, the LFA is a newly improved tool developed by EuropAid, in which the logic of a project is based on the results chain and the overall objective is to achieve impact in the long-term.

Specific objectives are reported below the overall objective and are composed of intermediary and final outcomes. Furthermore, indicators that measure the change produced by the intervention at every level must be disaggregated by gender.

The central columns in the table present three new elements: Baseline, Current Value, and Target/Beneficiaries. The Baseline is the data collected in the initial phase and includes starting values, while Current Value is the baseline data that is updated during project delivery and aims to update value on the basis of intermediary outcomes during the course of the project. As such, it requires a reference date.

With regard to the information required for the Target/Beneficiary column, it is important to note that it does not refer to project beneficiaries, but rather, to the final values of indicators that must be met (as stipulated in the project plan). In the remaining sections (activities, inputs, outputs, assumptions) the logical framework does not vary. The removal of the Expected Result row is among the largest innovations and there is now no longer an intermediary level that ties outputs to outcomes¹⁷.

The table below presents the *logframe matrix of a project*:

17 This change is in line with the "Horizon 2020" new Work packages approach. There continues to be a reference to results in the Results Oriented Monitoring (ROM) guide published by the European Commission in 2012 which still applies today.

FIG.8 - New logical framework

	Results chain	Indicators	Baseline (incl. reference year)	Current value (incl. reference date)	Target/ Beneficiaries (incl. reference year)	Sources and means of verification	Assumptions
Obiettivo Generale: Impatto	The broader, long-term change which will stem from the project and a number of interventions by other partners.	Measure the long-term change to which the project contributes. To be presented disaggregated by sex.	Ideally, to be drawn from the partner's strategy.		Ideally, to be drawn from the partner's strategy.	Ideally, to be drawn from the partner's strategy.	
Obiettivo/i Specifico/i: Outcome(s)	The direct effects of the project which will be obtained at medium term and which tend to focus on the changes in behavior resulting from project Outcome = Oc (possibly) intermediary Outcome = iOc	Measure the change in factors determining the outcome(s). To be presented disaggregated by sex	The starting value or current value of indicators.	The value of the indicator at the indicated date	The intended value of the indicators.	Sources of information and methods used to collect and report (including who and when/how frequently).	Factors outside project management's control that may impact on the outcome-impact linkage.
Output	The direct/tangible outputs (infrastructure, goods and services) delivered by the project. Output = Op Op 1.1. (related to Oc 1) Op 1.2. (related to Oc 1) (...) Op 2.1. (related to Oc 2) (...)	Measure the degree of delivery of the outputs. To be presented disaggregated by sex..	Idem as above for the corresponding indicators.		Idem as above for the corresponding indicators.	Idem as above for the corresponding indicators.	Factors outside project management's control that may impact on the output-outcome linkage.
Activities	What are the key activities to be carried out, to produce the outputs? (Group the activities by result and number them as follows: A 1.1.1. - "Title of activity " A 1.1.2. - "Title of activity" (related to Op 1.1.) A 1.2.1. - "Title of activity" (...) (related to Op 1.2.) A 2.1.2. - "Title of activity" (...) (related to Op 2.1.) (...)	Means: What are the means required to implement these activities, e. g. staff, equipment, training, studies, supplies, operational facilities, etc. Costs: What are the action costs? How are they classified? (Breakdown in the Budget for the Action)				Factors outside project management's control that may impact on the output-outcome linkage.	

With regard to the second approach, namely, the **Theory of Change (ToC) and the social value chain** (Figure x), **we have presented organizations with the necessary steps** to accurately develop their evaluation process in line with project planning.

FIG.9 - Processo

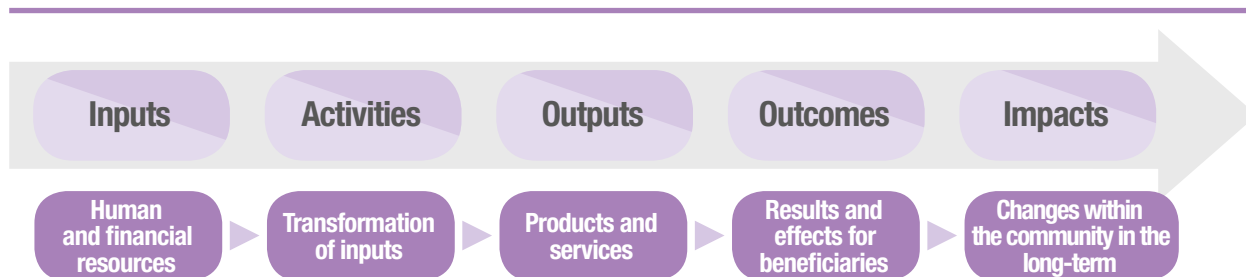


As part of the definition of the context of the analysis, the following elements must also be established: the evaluation objective and audience, its specific object (in our case, the organization itself), the resources and time available, and the individuals tasked with carrying it out.

The phase in which different actors are analyzed and engaged is a fundamental step that involves the identification and the mapping of all stakeholders, as well as determining their relative importance and securing their concrete engagement. As part of the evaluation process and in line with a participatory approach, different stakeholders (from beneficiaries and communities to partners co-delivering the project) should be actively involved in the initial feasibility assessment, as a regular part of the project cycle, in order to identify the evaluation methodology. In cases where stakeholders have a limited ability to jointly decide on the method to be adopted, engagement must take place during first phase of project implementation, during which a detailed assessment and project plan, as well as a possible baseline analysis, will be carried out. This is also the phase in which detailed indicators are selected, alongside the monitoring plan and tools which must be consistent with project information needs, aims and the chosen methodology.

In order to map the process of change, the logical framework elements must be defined and validated as a social value chain. These are presented in the figure below:

FIG. 10 - Elements of the logical framework/value chain



Once the chain has been constructed, the indicators linked to its different elements must be selected. The association of indicators is fundamental to be able to measure impact adequately. A correct methodology for the accurate development of indicators must include:

- A definition of the measurement methodology;
- The identification of indicators which can be *standard*, *ad hoc* or *proxies*. This report does not attempt to provide sector-specific indicators (KPIs), however, it draws attention to the existence of international frameworks of reference that can inform the selection of shared KPIs, such as, for example, the Global Impact Investing Network (GIIN) database¹⁸, Impact Reporting & Investment Standards (IRIS)¹⁹ and the Global Reporting Initiative (GRI);
- The selection of indicators that meet SMART criteria (Specific, Measurement, Attainable, Relevant, Time-bound);
- Data collection that takes place via tools such as questionnaires, interviews, shared databases chosen in line with the evaluation methodology that has been adopted;
- The development of indicators that include means of calculating quantitative indicators and means of describing qualitative indicators;

Impact evaluation is closely linked to the choice of measurement methodologies and by the definition espoused. In fact, social impact can be defined as:

18 The Global Impact Investing Network is a non-for-profit organization that aims to increase the scale and efficiency of impact investing globally. <https://thegiin.org/>.

19 The Impact Reporting and Investment Standards (IRIS) (<https://iris.thegiin.org/metrics>) is a catalogue of performance indicators, divided into sectors and operational areas, which can be used to evaluate social, environmental and economic impacts. IRIS is built on 40 metric taxonomies, making it adaptable to different sectors, geographic areas and performance types. It simplifies the task of comparing different investment opportunities in order to create performance ratings.

- The effects and changes produced by the activities carried out by an organization in the long-term for the community → qualitative methods.
- The portion of the total outcome that happened as a direct result of the intervention, minus what would have happened anyway in the intervention's absence → quantitative methods: counterfactual.

The first definition requires the transition from outcomes to impact to occur through qualitative methodologies, bringing outcomes into a long-term perspective that focuses on the community, rather than on direct beneficiaries. In certain specific cases, this transition may also be accompanied by the use of synthetic indicators, such as multipliers, which monetize the savings to public expenditure generated by the outcome.

The second definition requires outcomes to be calculated using counterfactual analysis. Here, impact is the result of outcomes defined through attribution, minus the effects that are not relevant or are deemed to be negative. Counterfactual analysis establishes: the effect of an intervention on an individual as the difference between the observed result with treatment and the result that would have been observed for the same individual without treatment. The analysis is based on the concept of a control sample which can be created using different methodologies that may include randomization techniques, *regression* discontinuity design or longitudinal analysis.

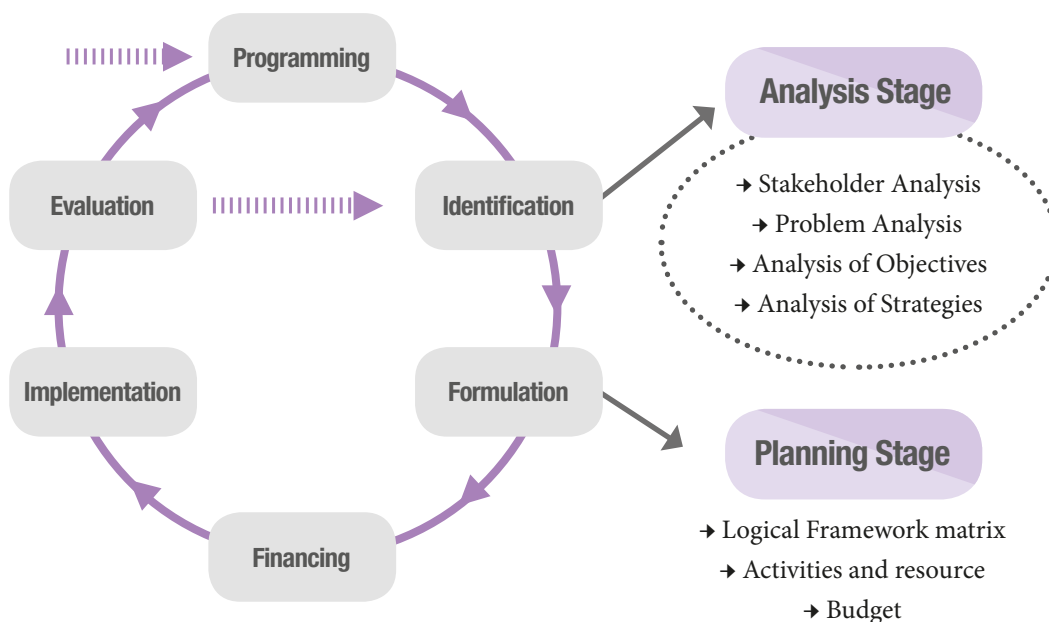
In both of these cases, the variable of “time” must be accounted for. In fact, while both analyses may be able to measure the level of outcome achieved, they will now allow for a complete measurement of impact, as this requires a time frame that spans beyond the project delivery period. This is why it is important to establish a process that allows specific concrete results to be achieved, beyond the end of the project itself.

4.4. Guidelines for the measurement of impact generated by development cooperation projects

The following paragraphs will attempt to highlight and guide the operational integration of both approaches, which appear to have a degree of logical continuity between them. First of all, there is absolute correspondence between the value chain of process-based models and the basis of the ToC approach.

A closer look at PCM and its conformity with the logical framework which lays out the necessary steps required to “complete” the matrix (*Figure X: Correspondence PCM & LFA*), reflects consistency with the Theory of Change (ToC) process described previously.

FIG. 11 - Correspondence PCM & LFA



In fact, the two approaches share several steps. When the logic of PCM is analysed together with logical framework and SIA (social impact assessment) and the Impact Navigator are selected as measurement methodologies, the following commonalities emerge:

STEP	PCM	LOGICAL FRAMEWORK	SIA	IMPACT NAVIGATOR
Defining the context of analysis	In the programming phase in order to establish the nature of the intervention, geographic area, sector and partners	-	First step in order to identify measurement aims	In the planning phase – needs and context analysis
Stakeholder analysis and engagement	In the identification phase through the analysis of the logical framework	Analysis phase in order to establish stakeholder needs and project aims	Second step in order to establish stakeholder needs and desired changes	In the planning phase – Establish project aims.
Defining the value chain	In the formulation phase through the LF planning phase	Planning phase, construction of matrix	Third step in the construction of the chain	In the planning phase – Development of logical framework
Defining the indicators to be linked to chain elements	In the formulation phase through the LF planning phase	Indicators must meet SMART criteria	Indicators must meet SMART criteria	Indicators must meet SMART criteria
Data collection and analysis for impact evaluation		Impact evaluation phase	In the evaluation phase and when estimates are being made	In the results analysis phase

Who should carry out social impact evaluation?

As a general principle, it is fundamental that project funders and leadership provide their full support to evaluation activities both formally and operationally: impact evaluation requires sufficient resources to be earmarked, specific procedures to be put in place and for responsibilities to be clearly delineated. Therefore, when the roles and responsibilities for carrying out social impact evaluation are assigned, the group of evaluators selected must possess the minimum skills required, which include:

- Experience and knowledge of the specific project sector and experience of evaluation within the project area;
- Good knowledge of the methodology and high quality personal standards;
- Objectivity (note that even an evaluator cannot be entirely objective because he or she will have options, values and will work on the basis of previous experiences. For this reason, the points in which the evaluator brings his or her own opinion into the evaluation must be clearly identified);
- Good communication skills (oral and written);
- Being honest and reliable as well as able to build relationships with beneficiary groups in a sensitive manner (for example, in relation to gender diversity).

Evaluations may be carried out by internal project workers or external service providers and this must be established on a case by case basis. Impact evaluation requires technical and methodological skills that are not available within an organization itself however, sometimes the organization would do well to develop them internally in order to increase its familiarity with the project and improve access to project information.

When evaluation takes a mixed form that includes internal and external elements, project staff members carry it out jointly with an external consultant who brings experiences and an external perspective to the process. This combines the advantages of both.

The following paragraphs describe the single steps of the impact evaluation and measurement process and serve as an operational guide for projects in the field of development cooperation. As previously mentioned, **this does not constitute a standardized model but rather, a process** that aims to integrate project planning activities which are integral to development cooperation, into evaluation. As a result, this report aims to reinterpret the process through a series of operational steps that can place the analysis of information regularly produced by a project within a perspective of impact.

4.5. Step 1 ■ Definition of the context of analysis

The projects that are developed and carried out should meet the needs of their target groups as much as possible. For this reason, if the definition of the context also includes impact, there can be an understanding of whether the project truly meets the needs of individuals. This will lead to more informed decisions being made around which project to continue delivering and which to discontinue, with a strong likelihood that delivery will lead to the planned effects of a project.

There are two ways to collect information for the purposes of defining the context and the needs analysis. It is often possible to obtain pre-existing data from reliable sources and previous assessments, however, if the information is unusable or insufficient, additional data must be collected and the volume will be determined by the human and economic resources available.

Therefore, the first important step is to understand the context of the project, namely, the primary social problem that it intends to address. Focusing on the specificities of the social problem will allow the target beneficiaries to be identified and as a result, uncover the impact that the project aims to produce.

CISP

In the case of the CISP project delivered in two Ethiopian regions in support of vulnerable children and adolescents, the context analysis was carried out through field investigations in order to obtain accurate prior knowledge of the territory; this involved direct consultation with privileged local witnesses and with local institutions at different levels. The benefits that can be derived from building direct relationships and collaborating with key local actors should not be underestimated. In this specific case, the fact that data was collected before the start of the project and through intensive collaboration with the districts and kebel, allowed good partnerships with local institutions to develop from the beginning. These became progressively strengthened until they were characterized by continued debate and exchanges of information. In any case, direct data collection can be supplemented by pre-existing documents to identify the potential actions of other actors that can be integrated and/or supported by one's own project. In this case, the aim was to reinforce the institutional network that protects vulnerable minors and strengthen the institutional and organizational capacity to fulfil institutional commitments.

Once the problem has been identified, the next step is to search for potential causes and effects. Here, it is useful to map the web of causes and effects starting from direct causes and effects and ending with the identification of the “causes of causes”. Subsequently, all these elements must be linked logically, paying careful attention to not leave any gaps.

CCM

The health project implemented by CCM (Comitato Collaborazione Medica-Committee for Medical Collaboration) in the Tigray region (Ethiopia) began with a context analysis and with the close and regular collaboration between the Italian and Ethiopian partners, who jointly identified the causes of the social problem being addressed and the barriers faced to its resolution. In this case, the primary barrier to improving mother-child health and achieving the development aim was the scarcity of human resources within the health sector. A more in-depth analysis uncovered the “causes of the causes” which, in this specific case, proved to be the lack of interest of qualified staff to stay and work in disadvantaged rural areas, namely, the so-called brain-drain in which professionals prefer to seek employment in countries or regions with higher GDPs. The causes were directly identified, making it possible to develop a more effective intervention strategy.

In order to understand the social problem and the context in which the project exists, including stakeholder needs and the environment in general, the following checklist can be completed:

Key question checklist for needs and context analysis within the framework of project impact evaluation:

- What is the social problem being addressed?
- What are the root causes of the problem?
- What are the effects of the problem?
- What is the scope and significance of the project at the local level?
- What specific situation does the beneficiary group face?
- What do the individuals within the beneficiary group need?
- What other services already exist within the chosen delivery location? What results have these services had?
- Are there any gaps in existing support programmes that must be filled?

4.6. Step 2 ■ Analysis of objectives and stakeholder engagement

Once the context and needs have been established, the next step in the impact measurement process is the definition of the project. In fact, only organizations that carefully consider and analyse their project objectives are truly able to evaluate the impact produced by their work. As previously mentioned, for the implementation of a development cooperation project to have a certain impact, it is fundamental that the primary stakeholders – namely, all the individuals, groups or institutions that are positively or negatively affected by the project and/or could influence it – are involved from

the project planning and objective-setting phases. It is equally important to regularly involve stakeholders across the process of social impact analysis. In this way, the quality of procedures and results can be ensured and objections can be avoided late into the project. A development cooperation project that aims to demonstrate impact can set itself different types of objectives, depending on the stakeholder type, which include:

- objectives that relate to the group of reference (outcome-objectives)
- objectives that relate to society (impact-objectives).

Outcome-objectives describe the benefits created by the project for the beneficiary group, as well as the changes that occur within the group as a result of the project. When an objective that relates to the target group is fulfilled, the result can and should be attributed (at least primarily) to the activities of the project itself.

COSV

The project implemented by COSV on the reproductive and sexual health of young people in Zimbabwe sets its aims in accordance with the Strategic National AIDS plan objective to reducing the HIV infection rate by 50%. In particular, the project aimed to **improve awareness, care, support and treatment services to young people**, in order to reduce Sexually Transmitted Infections (STI), as well as unwanted pregnancies among school girls and the subsequent school drop-out. The project made use of existing resources and established which activities were needed to fulfil set project objectives. Stakeholder involvement, as well as the impact on behaviour change they reported together with beneficiaries, emerged as particularly positive aspects of the project.

GVC

The involvement of direct beneficiaries and stakeholders is required in order to have a broader and more comprehensive view of the effectiveness of the project and identify both its objectives and potential contribution to social change. Migra-Safe is a project delivered by GVC in Cambodia aimed at promoting legal and safe migration. Stakeholders and direct beneficiaries were involved in setting project objectives, both in the planning phase, through focus groups, interviews and ad hoc questionnaires, as well as during delivery, through the detailed planning of aims and specific activities. This led to an evaluation that assessed both the intervention, as well as single activities.

Impact-objectives describe the changes generated at the societal level in the medium-to-long-term, or changes which the project contributes to create. In fact, in the majority of cases, effects of this kind are not exclusively attributable to project activities, but are also the result of factors external to the organisation. For this reason, it is advisable to define *contribution* at the project level, for example, by using the phrase “the project contributes to...”. The impact-objectives at the societal level to which the project contributes are expressed in more abstract terms compared to the objectives set

for the beneficiary group. Rather than being connected to the members of a group, the former are tied to society in general, or at least to a part of society, for example, to the population of a specific region. As well as depending on different factors, desired changes are often observable in the long-term. In some cases it is therefore unrealistic to assume that these changes can be achieved within the delivery time frame of the project.

COSV

The MusicBridge project delivered by COSV focused on three different outcome areas: institutional, creative production processes and community. In particular, the project was found to have:

- Significantly strengthened the competencies of local partners in relation to the technical and administrative management of international cultural exchanges. The primary outcome identified was the creation of a network of individuals and institutions from the Pacific, Southern Africa and Europe.
- Expanded the horizons and perspectives of the musicians involved, increasing their awareness of being artists and of the opportunities available to pursue a career within the field.
- Created income opportunities for the Espiritu Santo communities in Vanuatu.

In this case, the final evaluation assessed the contribution of COSV, the lessons learned and emerging good practices, the potential of the ACP-EU network within creative industries, the project's sustainability and provided recommendations to increase impact in the long-term.

Below are suggestions for setting objectives, which should:

- Be defined as if results have already been achieved;
- Be framed in a positive way and describe the actual situation desired;
- Avoid negations which place unnecessary attention on problematic situations;
- Specify the project's direct recipient groups who benefit from the changes and use their voice to describe the ways in which their concrete living conditions change once project outcomes are achieved;
- Set impact-objectives at the societal level, in order to transform social problems into clear statements, as if the problem no longer existed;

CISP

For the project delivered by CISP in Armenia that aimed to maximise the socio-economic impact of migration, the impact analysis carried out was guided by two key questions: is the project able to improve/make a difference to the lives of beneficiaries? Is the project making headway in shaping the definition of a good migration and development policy and an action plan relating to these two issues? In this case, the objectives identified and connected to the strategies and policies on immigration and development in Armenia are essentially: to maximize the impact of migration by involving the diaspora and migrants in development activities through remittances and the competencies and know how of individuals who live or have lived abroad; to prevent illegal immigration through the creation of economic opportunities or alternatives for individuals at risk; and as a final objective, in light of the fact that it is a geographical area that hosts refugees and asylum-seekers from Siria, to provide immediate assistance and promote socio-economic integration. The identification of these objectives allowed for a clear understanding of the tools required to promote change, while taking into account the actions and presence of other important local actors to collaborate and exchange views with.

GVC

The Make a Fruit Fair! Project delivered by GVC adopted the Outcome Harvesting approach which is based on the idea that an outcome is generated when someone does something differently. For this reason, impact was identified via the question: “what is different? In particular, the analysis required was based on a qualitative evaluation composed of the following two steps:

- Identifying the desired impact based on project objectives, intermediary and long-term outcomes and on the change mapped and discussed with partners.
- For each outcome, evaluators asked “*who is doing what differently?*”, while thinking of the the relevant stakeholders as well as the planned project results. Here, specific questions were asked firstly, to uncover the project’s contribution to each of the changes listed and identify other factors that may have helped produce the change, and secondly, to find evidence to support these claims. Finally, the project’s contribution was assessed in order to determine what could have been done differently and the actions required to improve the intervention.

Checklist for the formulation of project objectives

- What effects should project activities have on the beneficiary group?
- How will social, financial or living conditions change in general?
- What would the situation be for society in the future (positive) if the problem no longer existed?
- What is the time frame within which the objective must be met?

4.7. Step 3 ■ Defining the value chain – logical framework

After having examined the social problems and corresponding needs of the beneficiary group, the organisation must define the objectives of the intervention and the value chain through the project's logical framework, in order to measure impact. Some inputs (resources) must be available in order to fulfill objectives and produce outputs (services and products). If outputs are used by the individuals who make up the beneficiary group, the desired changes for the beneficiary group (outcomes) can be achieved, which, in turn, can lead to changes at the societal level (impact). This leads to a change in social conditions which can give rise to new needs and may require project objectives to be readapted and resources and project outputs to be reassessed. Planning, learning and adaptation all form part of a continual process of the measurement, evaluation and reformulation of objectives that takes place during and beyond the project's delivery time frame.

Interim evaluation

In the adaptive evaluation approach, evaluation is placed within the cyclical perspective of an intervention (planning-learning-adapting) and the choice to conduct interim evaluations is an important step that allows the actors involved to reflect methodically on the strategy adopted and the interim results achieved, in order to better inform subsequent steps.

Interim evaluation can be carried out by a consultant (or a team of consultants) internal to the organisation (or a consortium of partners) that is not directly involved in project implementation but knows the intervention and is well-acquainted with the objectives set by the organisation (or consortium). It holds the advantage of being able to address aspects related to management in some depth, and the *neutral* nature of its enquiry encourages staff and stakeholders to cooperate in order to identify recommendations that can effectively be applied. Conversely, the more one-sided nature of the intervention can be counterbalanced by the use of external consultants in select phases of project evaluation. The evaluation methodology initially selected will be also adopted in the internal interim evaluation, allowing for an analysis that is more comparable with other evaluations that may take place during the life of the project.

GVC

NelFor GVC's Migra-Safe project aimed at promoting safe migration and combating against human exploitation and trafficking in Cambodia, a mid-term evaluation was conducted to analyse the results achieved in relation to pre-set criteria and objectives, as well as to identify potential gaps or omissions. Here, the involvement of stakeholders and direct beneficiaries was given particular importance as these groups provided information that subsequently informed recommendations to improve project activities, both during delivery, as well for the future. Some of these pointed to the need to conduct lobbying activities, to seek greater involvement from anti-trafficking national and provincial taskforces, and to introduce training on specific activities such as agriculture. As such, stakeholder involvement led to a broader and more comprehensive analysis of the effectiveness of the project: 80% of suggestions made were included in the activities and projects subsequently designed. Furthermore, the European Delegation conducted a ROM review and assessed the project using four criteria – relevance, efficiency, effectiveness and sustainability; the emerging report contained several recommendations that were taken into account and led modifications to be made to the project, on the basis of recommendations contained in the mid-term evaluation. This made the intervention increasingly effectual.

In order to develop the logical framework previously described, it is important to know its various elements and the relationship between them.

Inputs

Inputs are the resources required for the implementation of the project. They first of all include staff, volunteers and their hours of work, followed by financial resources, building costs and the cost of equipment required for project implementation. In order to create a feasible project plan that includes impact, organizations should make a list of all necessary resources.

Outputs

Outputs are the (quantifiable) products and services generated by a project, namely, what a project does and offers, as well as its effective use by a beneficiary group. Outputs can be directly related to inputs to allow conclusions to be drawn regarding the efficiency of project implementation and to determine, for example, the participant unit cost for the project.

Outcomes

The transition from outputs to outcomes and impact is crucial to the analysis of the project's impact. While outputs are required for the project to meet its desired outcomes, outcomes are only

generated when beneficiaries effectively experience positive changes. For many years, project planning and evaluation carried out with PCM focused on project activities and outputs, essentially, on “what happens during the project” and on “what the project achieves”. This partly explains why organisations tend to place more importance on the description of activities and outputs and pay less attention to the connections between these two elements and outcomes and impact. As a result, to date, many evaluations simply describe the nature and scale of project activities, providing information that is confined to the sphere of outputs, such as, for example, the number of users and beneficiaries engaged. In order to conduct an evaluation that goes beyond the direct results of the project and present findings on impact, the focus must be placed on the change generated by the project, even though obtaining concrete information that points to its actual impact is not an easy task.

Outcomes for project beneficiaries are composed of the following sequential steps:

1. Changes in knowledge, awareness and abilities;
2. Changes in behaviour or actions;
3. Changes in living conditions (financial and/or social conditions).

This distinction is useful as the changes reported in participant knowledge, behaviour and actions are prior conditions required to meet the objective of improving the living conditions of participants.

Nonetheless, it is often the more concrete and measurable outcome that relates to the change in beneficiary living conditions which must be reported in project evaluations (a young person who has found employment or the increased earnings of an entrepreneur). While there is nothing wrong with this trend as such, the intermediary changes that must be achieved in order to meet a set outcome are at least *as* important as the outcome itself, thus, the focus should not solely be on the type of outcome outlined above, but also, on intermediary results.

Impact

While the results achieved for the beneficiary group are called outcomes, impact describes the changes produced at the societal level. These can include changes to socio-economic conditions. Nonetheless, in light of the fact that, in the majority of cases, it is neither useful nor possible to assess the changes that occur for society in general, impact commonly refers to a specific part of society, for example, a population within a set geographic area in which the project is being delivered.

It is important to note that changes do not occur from one day to the next, but rather, are part of a step-by-step process. Impact only occurs once a series of prior changes have taken place: firstly,

there must be a change in knowledge among beneficiaries, secondly, beneficiaries must use their newly-acquired knowledge. Finally, this leads to a change in living conditions after which the overall objective of impact can be met in relation to the target population.

If the results only emerge after some time

Very often, the effects of projects in the field of development cooperation will only emerge after a period of time. In this case, conducting an evaluation requires data to be collected after the project has ended. For this purpose, there must be a up-to-date contact database and participants must be notified that they will be contacted after the project has ended. An additional challenge lies in the fact that, after the end of the project, beneficiaries may be subject to other factors that can influence the effects that are rightfully attributed to the project. In order to deal with this situation, the logical framework is useful as a step-by-step guide to assess the gradual achievement of outcomes together with the verifiable results achieved during the course of the project. If, for the duration of the project, changes in the awareness, knowledge, abilities and levels of behaviour have been reported, one can assume that the project will also have long-term effects on the living conditions of beneficiaries.

Logical framework development

The logical framework can be developed in two ways: either by starting with desired impact and ending with the necessary inputs or by beginning with inputs and ending with impact.

CCM

In the case of CCM's project in Tigray (Ethiopia) the first step was to identify the social problem to be addressed by the project, its root causes, as well as the local actors who were actively seeking a solution to the problem.

The general objective was to improve mother-child health and more specifically, to strengthen the health system by supporting the regional health care system to train staff that would then be placed in first-level facilities. The development of a detailed and comprehensive logical framework required the project's expected results to be identified in line with set objectives, as well as the activities required. Furthermore, in accordance with logical framework specifications, external conditions and objectively verifiable indicators tied to each level of analysis were reported.

In order to develop a logical framework that begins with impact and ends with inputs, the starting point should be impact at the societal level, and outcomes for the beneficiary group. Subsequently, for each stage, the following questions and answers apply:

IMPACT → OUTCOMES	What must happen to allow socio-economic conditions at the societal level to improve?	There must be changes reported in the living conditions of the beneficiary group!
OUTCOMES	What must happen in order to generate changes in the living conditions of beneficiaries?	There must be a change in the behaviour or actions of individuals within the beneficiary group!
OUTCOMES	What must happen in order to generate changes in the behaviour and actions of beneficiaries?	There must be a change in certain aspects that relate to the knowledge and attitudes of individuals within the beneficiary group!
OUTCOMES → OUTPUTS	What must happen in order to generate changes in the knowledge and attitudes of beneficiaries?	Participants must use the service!
OUTPUTS	What conditions must be met to ensure that services are readily accessed?	Services must be available!
OUTPUTS → INPUTS	What conditions must be met to ensure that services are available?	Resources must be sufficient!

Alternatively, the development of a logical framework that begins with inputs and ends with impact requires an analysis of the “if-then” connection between the different elements. In this case, the steps are as follows:

	IF...	THEN...
INPUTS → OUTPUTS	IF the input resources required are available and have been invested in the program,	THEN the planned activities or services can be delivered.
OUTPUTS	IF the products and services are available,	THEN they can be used by the beneficiary group.
OUTPUTS → OUTCOMES	IF the products and services are used or accessed,	THEN the conditions to achieve desired changes in the knowledge and attitudes of individuals within the beneficiary group have been met.
OUTCOMES	IF the desired changes in the knowledge and attitudes of the beneficiary group are achieved,	THEN the beneficiary group can act in new/different ways.
OUTCOMES	IF beneficiaries act in new/different ways,	THEN changes in living conditions of beneficiaries can occur.
OUTCOMES → IMPACT	IF the living conditions of project participants change,	THEN the new situation contributes to creating changes at the societal level.

ELIS

The Marcona Project delivered by ELIS in Perú is an intervention that aimed to improve the socio-economic conditions of the community. Firstly, stakeholders and their expectations of potential changes were identified. This was followed by the assessment of project inputs, outputs and outcomes. Outcomes emerged from an analysis of the baseline and through information obtained via stakeholder involvement. For each outcome and every corresponding project output, a length of time and appropriate indicators were selected. All the outcomes reported by stakeholders in the consultation phase are part of the theory of change.

Esempio

Main beneficiaries:	Output:	Outcome:	Indicator:	Source:	Quantity and time:
Fishermen of 16 local OSPAs (COPMAR)	137 fishermen trained on the topic “ <i>Safe diving and safety at sea</i> ”	Achieved diving licenses and reduced accidents caused by diving dangerous practices (Activity 1)	Reduction in number of accidents at sea (90%)	Baseline interviews; monitoring report data; local OSPAs registers....;	4,5 (90% of 5 accidents per year)

Checklist for completing the logical framework

- Which of the necessary inputs are available and were invested in the planning of the activity and the delivery of the services offered?
- What products and services are available and can be used by the beneficiary group?
- What desired changes in the knowledge and attitudes of individuals within the beneficiary group were achieved?
- What new/different actions were pursued by the beneficiary group?
- What changes in the living conditions of participants were made possible?
- What is the new situation that contributed to creating change at the societal level?

4.8. Step 4 ■ Indicator Step

The following step relates to the development of indicators in order to verify whether project objectives have been met and changes have been generated, for the purposes of impact measurement.

Direct and indirect indicators

Direct indicators are often created on the basis of project objectives and are directly linked to the outputs and outcomes that the evaluation aims to describe and measure. A classic example is the indicator for the increased income of project beneficiaries following the delivery of training activities that aim to place beneficiaries into work. However, it is not always possible to build such indicators and in order to remedy the problem, indirect indicators or approximations of observed

or desired change can be used. For example, the number of people who access a specific free meal service can be used to estimate the poverty level within a geographic area. In particular, indirect indicators are useful to describe the qualitative changes experienced by beneficiaries which may be difficult to measure, such as the quality of life, awareness or motivation levels.

Indicators are required at all levels of the logical framework in order to establish whether or not the project is following its planned course. Indicators are therefore necessary as reference points in the the process of evaluating and measuring project impact.

Input indicators

Input indicators are important for two main reasons: firstly, they provide information surrounding the resources invested in a project and secondly, they allow the project's efficiency to be evaluated. If inputs are compared to outputs, outcomes and impact, the level of inputs required to produce a given level of outputs (efficiency) or outcomes and/or impact (social return on investment) can be established.

Output indicators

As previously mentioned, outputs are the basis of the corresponding impact of a project and a pre-requisite for its achievement. During the initial project implementation phases in particular, outputs may be the only elements that can be monitored, in so far as changes and impact are only generated in subsequent phases. Therefore, even if the changes produced cannot yet be verified, once activities are delivered, initial output indicators should be easily measurable.

Outcome and Impact indicators

Finally, outcome and impact indicators are required in order to evaluate whether the project has generated changes and impact and to what extent it has done so. While impact indicators measure results and social change in the long-term, it is important to create outcome indicators for the different levels of the logical framework, in order to understand and illustrate how project objectives have been achieved. In some cases, a single indicator is sufficient to measure project objectives. This applies in particular, to quantitative data, for example, in the measurement of the number of participants that attend a training course who go subsequently enter the workforce. However, at times, there are more complex exploratory objectives which require several indicators that are usually a combination of qualitative and quantitative factors aimed at measuring the objectives or changes achieved. In fact, often, the underlying aim of the participatory approach adopted by

development cooperation projects and required for the purposes of internal development, is the active engagement of the community. Methods and indicators that can capture the participatory component are essential in order to highlight results connected to community response mechanisms that are triggered. On occasion, the participatory approach can lead initial objectives to be redefined. In these cases, participation can be a project result in itself, in light of its potential effect on the socio-economic advancement of the community.

COSV

For COSV's project in Zimbabwe on sexual and reproductive health, indicators were identified through documents (project proposal, baseline, project reports, European Union ROM review, interim evaluation) and benchmarks. Both qualitative and quantitative indicators were adopted, which the team considered to be appropriate for the measurement of the project. Thanks to the recommendations that emerged from the ROM review, the five chosen indicators (reduction in the local rate of teenage pregnancies, early school drop-out rates, STIs prevalence among young people and the increase in the use of contraception and in circumcisions among young men) were reviewed and subsequently reformulated to be more specific, on the basis of suggestions made (for example, by inserting the terms 'voluntary' and 'medical' in relation to male circumcision) to allow change to be measured more precisely.

The indicator selected must be deemed significant and measurable. To this end, it should meet all the SMART criteria previously mentioned (specific, measurable, accepted, realistic, time-bound). Furthermore, it may be appropriate to include other criteria like the geographic location (for example the city or district).

Below is some advice for the creation of indicators:

- Formulate indicators in a way that allow its objective, audience and time frame to be clearly stated;
- Consider the appropriate form or unit of measurement of an indicator (quantity, total, average, percentage of total, sum, change in percentage etc.).

In order to ensure that an indicator is practicable, evaluators must consider the feasibility of collecting the necessary data required for measurement. They must verify whether the data source is appropriate and accessible and whether the effort required will match the benefits derived from the information that is obtained.

CUAMM - The evaluation of hospitals

Evaluating hospital performance in development cooperation programmes is not easy, for example, in cases where the intervention is in favour of hospital facilities in the non-for-profit sector and where governance systems, funding and human resource management systems are very different to those in the public sector. There are international evaluation models of hospital performance which generally focus on improving the quality of services delivered and can act as a reference point to set an accreditation system in motion. It must be pointed out that there is an overall limit common to all the approaches to the evaluation of hospital performance adopted in high-resource countries (eg. *DRG*, *Case Mixed and Severity of Illness Index*, etc.), namely, that they are not easily adapted to low-resource contexts in which data collection and analysis is very problematic. A simple but valid measurement tool is the *Standardized Unit of Output* (SUOP). This is a complex indicator that brings together outpatient visits, hospital stays, prenatal visits, births and vaccinations in a system that assesses the use, efficiency, productivity and equity of hospital performance, via the assignment of different weights. This method allows for the various dimensions of performance to be considered in conjunction with one another and to be monitored in timely way, while also providing essential help to hospital managers.

4.9. Step 5 ■ Data collection and analysis for impact evaluation

If pre-existing data are insufficient or inexistent, new data must be collected. There are several ways to establish the most reasonable and practical data collection method. This may depend on the quantity of resources available but also on which data is required, its level of depth and what the information needs are.

Data collection methods

An appropriate data collection method must be selected on the basis of the project's specific lines of enquiry and the corresponding indicators. In this light, the distinction should be made between quantitative and qualitative data collection methods.

Quantitative methods are used if the information required can be expressed as numbers or as ordinal values and if it includes a measure, count and quantitative analysis of secondary data (for example, through the use of statistics), different forms of questionnaires and tests, as well as structured observations. Quantitative data are particularly useful in cases where:

- A general quantitative picture is required;
- Results must be presented in an accurate way;

- Comparisons will be made between groups and/or people;
- Assumptions surrounding statistical relations between problems and causes have been made and must be tested.

Qualitative methods allow qualitative data to be collected. Qualitative data cannot take the simple form of numbers. They have a predominantly descriptive role and lead to a more in-depth understanding of situation and context. Qualitative statements are crucial to the management of the project and to evaluate impact and the causal relationships and links identified are as useful as quantitative statements. In this case, qualitative data collection methods are marked by a focus on the “hows” and on “whys” of a specific situation or context. Qualitative methods are particularly appropriate in the following situations:

- The situation within a given context requires more in-depth examination;
- The objective is to uncover the perception of individuals or groups in relation to their situation or to reveal their expectations and desires;
- It is necessary to prove and keep track of the development of the changes generated.

Qualitative methods include various forms of interview (individual and focus groups), observations and document analysis, among others.

The choice of the data collection methods to adopt depends on the information required and what its use will be. For example, funders and donors usually receive quantitative data and a summary of qualitative results. The latter are more commonly used in situations where a more in-depth understanding of the project is required or to describe the qualitative change generated by a project. As already mentioned on several occasions in this report, when the right project and contextual conditions exist, an evaluation that seeks to provide a comprehensive picture should ideally employ quantitative *and* qualitative methods, making use of both synthetic and in-depth measures.

In general, the descriptions of project evaluations considered in this report have revealed that both qualitative and quantitative data and indicators are used in evaluation practice. The data collection methods adopted in order to create baselines and carry out monitoring and evaluation activities range from interviews, focus groups, questionnaires, videos, observations, testimonies and quotes, and the use of pre-existing data from official documents. These different types of data are collected through mathematical-statistical tools in the case of quantitative data or used in story-telling techniques in order to strengthen the understanding of the change generated.

FIG.12



Source: Social Impact Navigator, Phineo

Data collection allows for the assignment of values, which is required for the measurement of resources, results and the changes generated by the project.

Basic values

Data that relates to basic values contains information on the baseline prior to project commencement. If this information is missing, the extent to which changes occur from the beginning of the project cannot be measured.

COSV

For COSV's project in Zimbabwe, two baselines were collected with the support of local COSV partners. These allowed for a more in-depth analysis of certain aspects connected to the social problem identified. In particular, the data collection allowed project plans to be defined, by assessing: access to SRH services, young people's knowledge of transmitted diseases, the social backgrounds of young people and the influence on their sexual/reproductive behaviour and habits. Thanks to the data collection at the beginning of the project, the ROM review and monitoring activities, a final project evaluation was successfully carried out.

Aim values

In order to be able to make claims about the achievement of objectives, it is necessary to establish criteria in order to identify whether indicator values reflect the fulfilment of set aims.

In order to identify the values connected to these aims, realistic expectations must be set on the basis of personal experience or on the experiences of organisations delivering similar projects. Once the data that measure changes generated by the project have been collected, different types of comparisons can be made, on the basis of knowledge purposes:

Data comparison type	Mode
Before –after comparison - Longitudinal analysis	The comparison of data collected before and after the intervention allows the changes that occurred in time to be measured.
Comparison with the value aim	The comparison of data that relates to results and to the effective value aim set for the project.
Comparison between the different configurations of the project - Sensitivity	The comparison between the different data configurations of a given project allows contributing factors to be identified and an understanding of how they contribute to the change generated.
Comparison between beneficiary and non-beneficiary groups and sub-groups - Counterfactual	The comparison between beneficiaries and non-beneficiaries allows the changes produced by the project to be measured.
Comparison between projects and benchmarking	The comparison between data collected for similar projects by different organizations or similar projects in different locations.

GVC

In the case of GVC's project in Cambodia which aimed specifically to increase the awareness of risks associated with illegal migration among the target group, a baseline was created to identify initial levels of knowledge regarding migration processes and the various implications connected to illegal and legal channels. The data gathered provided starting values for the indicators tied to objectives. These were then compared to the results derived from a variegated information program (which included peer to peer activities, a communication campaign delivered through theatre and other art forms). Data were collected via questionnaires that were disseminated to a representative sample.

When the results obtained are not satisfactory

On occasion, the results obtained may suggest that project objectives have not entirely been met. In these cases, the organization should take the evaluation results into account and use them as a basis for critical reflection and learning purposes. For this reason, it is important to understand the reasons behind specific discrepancies, which could be internal or external to the organization. It is equally important to continue to monitor the project throughout its stages of delivery, to avoid any final surprises. In any case, organizations should be aware of any differences that emerge in relation to initial expectations and communicate them to stakeholders in a transparent manner, informing them of the measures that the organization intends to take, as part of a continual process of learning and improvement.

4.10. Step 6 ■ Communicating results

It is important to share the work conducted with all the individuals who take part/have taken part in the project for training and learning purposes. As previously mentioned, a culture of impact should be nurtured in which the evaluation process becomes integral to the work conducted by an organization, from the project planning phase, to the delivery phase and beyond.

One of the main tools to communicate results and share information on the evaluation process is the Impact report which can be used for both internal and external communication with the stakeholders identified at the beginning of the process.

Finally, on the basis of the evaluation audience, communication tools other than the report, such as seminars held at local organizations, social networks and webinars, are particularly suited to creating moments of exchange between the various project teams based in different countries.

Of particular note are the recent tools employed in adaptive evaluation processes which allow for the real-time socialization of results via web platforms, through which information can be shared among the different actors involved. These tools provide information to decision-makers who can take potential remedial action, to direct and indirect beneficiaries who can monitor and actively take part in the delivery of activities, to donors who can be kept up-to-date on the progress of initiatives, and finally, to organizations who can be transparent and obtain broader consensus on initiatives underway, with a view to encouraging the direct engagement of interested communities and potential donors and investors.

Appendix 1.

Methods for the quantification of impact (Puri et al., 2015)

Methods	Description	Pros	Cons
EXPERIMENTAL DESIGN			
Randomised controlled trial (RCT)	A sample of eligible subjects is randomly assigned into those who receive the intervention and those who do not. Impact is the difference in outcome between the two groups.	- Straight forward estimation (difference in means).	- Requires a comparison group; - Requires check of balance (i.e. whether randomization was successful). If randomization is not successful, then the results are not valid.
QUASI-EXPERIMENTAL DESIGN			
Difference-in-difference	Outcome of programme beneficiaries and nonbeneficiaries are compared before and after the intervention. The relative change in outcome is the impact of the programme.	- This approach deals with the problem of unobservable differences between treatment and comparison groups.	- Requires baseline data; - Requires comparison group; - Responsibility of ensuring balance in levels and trends is on the research team and usually requires a lot of data to ensure.
Regression discontinuity	A cut-off determines who is eligible to participate. Outcome of beneficiaries and nonbeneficiaries close to the cut-off line are compared.	- Does not require baseline data although it's desirable to have it.	- Requires comparison group; - Requires establishing that the comparison group at the cut-off is similar to the treatment group.
Matching	Programme beneficiaries are compared to a group of nonbeneficiaries that is constructed by finding people whose observable characteristics are similar to those of the people in the treatment group.	Does not require baseline data, except "matching variables" (that can be obtained from secondary data sources such as RLMS, DHS, etc.).	- Requires a comparison group; - Requires data on matching variables; - Assumes there are no differences in unobservables.
Instrumental variables	Participation in a programme can be predicted by an incidental factor, or instrumental variable, that is uncorrelated with the outcome (other than by predicting participation).	- Does not require baseline data; - The counterfactual is determined by the programme.	- Requires strong assumption that the instrument affects the outcome only through one specific channel that affects selection but does not directly affect the outcome.
Case-control (from medical studies)	An observational study is one in which subjects are not randomized to the exposed or unexposed groups; rather, the subjects are observed in order to determine both their exposure and their outcome status and the exposure status is thus not determined by the researcher.	- Can be less costly than RCTs: no randomization involved and fewer subjects observed; - Similar to matching: do not require baseline data, except data on "matching variables".	- Requires comparison group; - May require more waves of follow-up data (tracking mechanism); - The burden of proof for showing that the subjects are comparable and that no other reason may have brought about the observed change in outcome

Appendix 2.

Examples of Mixed Method Evaluation (Bamberger, 2012)

TITLE	TITLE AGENCIES SUPPORTING THE PROGRAM AND CONDUCTING OR COMMISSIONING THE EVALUATION
A. Quantitatively oriented evaluation designs	
1. Post-conflict reconstruction in Liberia.	DFID and the International Rescue Committee
2. Long-term evaluation of the Tostan program to reduce female circumcision in villages in Senegal.	UNICEF
3. Evaluating a conditional cash transfer program in Kazakstan.	Save the Children
4. Impact evaluation of FAO emergency and rehabilitation work in rural DRC.	FAO
B. Qualitatively oriented evaluation designs	
5. Evaluating the impacts of a gender-based violence prevention program in El Salvador.	Oxfam America
6. Life and livelihoods food security program in Bangladesh.	USAID, Save the Children and TANGO International
7. Evaluation of the UNICEF Education Programme in Timor-L'Este	UNICEF
8. Evaluating the equity-outcome of the Nepal Education for All Project.	UNICEF, NORAD and other partners
9. Evaluating the equity outcome of the Cambodia Community-Led Total Sanitation Project.	UNICEF
10. Inter-Agency Real-Time Evaluation of the Humanitarian Response to Pakistan's 2009 Displacement Crisis. ADB	UNICEF and various partners
11. Evaluating the road construction component of the Eritrea Com-munity Development Fund.	World Bank
12. Evaluation of the Egyptian Community Schools Project	UNICEF
13. Evaluation of the Tanzania Community Justice Facilitation Project.	UNICEF
14. Evaluating UNICEF's Response in the area of Child Protection in Indonesia, to the 2004 Indian Ocean Tsunami.	UNICEF
C. Balanced evaluation design giving equal weight to QUANT and QUAL approaches	
15. Evaluating the Kecamatan Development Project in Indonesia.	World Bank and Government of Indonesia
16. Evaluating the Indian Panchayat Reform Program	World Bank
D. Meta-analysis (secondary analysis of a number of evaluations to identify general findings)	
17. CARE International. Impact Evaluation Report 2005–2010. Latin America and the Caribbean.	CARE International

BIBLIOGRAPHY

AECID (2013) *Spanish Cooperation Evaluation Policy*

<http://www.cooperacionspanola.es/en/evaluation-publications>

ADB (Asian Development Bank), (2006), *Impact Evaluation Methodological and Operational Issues*

<https://www.adb.org/sites/default/files/institutional-document/33014/files/impact-analysis-handbook.pdf>

AFD (2013), *AFD's Evaluation Policy*

http://www.afd.fr/webdav/shared/RECHERCHE/Evaluation/document/AFD_evaluation_policy.pdf

Ashoka (2013), *How Do You Know When You've Revolutionized An Industry? Ashoka's Approach To Assessing Impact*

<https://www.ashoka.org/files/2013-Impact-Study-FINAL-web.pdf>

Bamberger, M. (2012). *Introduction to mixed methods in impact evaluation. Impact Evaluation Notes*, 3, 1-38.

Beach, D. and Pedersen, R.B. (2013) *Process-Tracing Methods: Foundations and Guidelines*, Ann Arbor MI: University of Michigan Press

Bengo, I., Arena, M., Azzone, G., Calderini, M. (2016). *Indicators and metrics for social business: a review of current approaches*. *Journal of Social Entrepreneurship*, 1(2): 1-24.

Burdge R. J. (2003), *The practice of social impact assessment background*. *Impact Assessment and Project Appraisal*, 21(2): 84-88.

Cameron, D.B., Mishra A., Brown A. N. (2016). *The growth of impact evaluation for international development: how much have we learned?* *Journal of Development Effectiveness*. 8(1): 1-21.

Center for Global Development (2006), *When will we ever learn?*

<https://www.cgdev.org/publication/when-will-we-ever-learn-improving-lives-through-impact-evaluation>

Costa, E., Pesci, C. (2016). *Social impact measurement: why do stakeholder matter?* *Sustainability Accounting, Management and Policy Journal*, 7(1), 99-124.

Department for International Development (2013), *International Development Evaluation Policy*.

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/204119/DFID-Evaluation-Policy-2013.pdf

DEVCO (2016), *PRAG Procurement and Grant for European Union External Action – Practical Guide*

Development Trust Association (2008), *Fit For Purpose: the Development Trusts Association healthcheck for community enterprise organisations*.

<http://locality.org.uk/wp-content/uploads/Fit-for-Purpose.pdf>

Directorate General External Relations, Directorate General Development, EuropeAid Co-operation Staff Working Document (2015) *COMMISSION STAFF WORKING DOCUMENT Launching the EU International Cooperation and Development Results Framework*, EC

Duflo E. Jameel A. L., Crépon B., Parienté W., Devoto F. (2008), *Poverty, Access to Credit and the Determinants of Participation in a New Micro-credit Program in Rural Areas of Morocco*, Imapct Analysis Series, N°2

<https://www.oecd.org/countries/morocco/42022186.pdf>

Earl S., Carden F., Smutylo T. (2001), *Outcome Mapping. Building Learning and Reflection into Development Programs*. International Development Research Centre

http://www.outcomemapping.ca/download/OM_English_final.pdf

Ebrahim, A. S., Rangan, V. K. (2010). *The limits of nonprofit impact: A contingency framework for measuring social performance*. Harvard Business Review Online

EC (2011), *Evaluation Matters. The evaluation policy for European Union International Development Co-operation*

https://ec.europa.eu/europeaid/evaluation-matters-evaluation-policy-european-union-development-co-operation-0_en

Elliston K., Maconachie M. (2002), *Morice Town Home Zone. A prospective Health Impact Assessment*

<http://www.who.int/hia/examples/house/whohia061/en/>

Epstein M. and K. Yuthas. (2014). *Measuring and Improving Social Impacts: A Guide for Nonprofits, Companies, and Impact Investors*. Berrett-Koehler Publishers

EU International Cooperation and Development, *First Report on selected result*, July 2013 – June 2014 (EU, 2016)

OECD DAC, (2012) *Evaluating Budget Support, Methodological Approach*,

G8 - Social Impact Investment Taskforce (2014). *Measuring Impact*. Subject paper of the Impact Measurement Working Group

<http://www.socialimpactinvestment.org/reports/Measuring%20Impact%20WG%20paper%20FINAL.pdf>

Gertler, P. J., Martinez S., Premand P., Rawlings L. B., Vermeersch C.M. J..(2016), *Impact Evaluation in Practice*, Second Edition. Washington, DC: Inter-American Development Bank and World Bank

<https://openknowledge.worldbank.org/handle/10986/25030>

Global Social Venture Competition, GSV, 2012. *Social Impact Assessment Guidelines*

http://www.i-edu.org.cn/gsvc/download/2011_GSVC_SIA_Guidelines.pdf

Grabenwarter, U., and Liechtenstein, H. 2011. *In Search of Gamma-An Unconventional Perspective on Impact Investing*. IESE Business School, University of Navarra.

Grieco, C. (2015) *Assessing Social Impact of Social Enterprises: Does One Size Really Fit All?* Springer International Publishing.

Hearn S. (2016), *Outcome Mapping and empowerment: the experience of SAHA in Madagascar*

<http://www.outcomemapping.ca/resource/outcome-mapping-and-empowerment-the-experience-of-saha-in-madagascar>

Hornsby, A. (2012). *The Good Analyst. Impact Measurement and Analysis in the Social-Purpose Universe*. Investing for

Human Foundation (2015), *Presentazione del rapporto finale di valutazione SROI*

<http://www.humanfoundation.it/ita/il-nostro-lavoro/materiali/24-sroi-padova/file.html>

International Initiative for Impact Evaluation (3IE), (2011). *3IE Impact Evaluation Practice A Guide for Grantees*

<http://www.3ieimpact.org/strategy/pdfs/3ie%20impact%20evaluation%20practice.pdf>

IOB Evaluation (2012), *Turning a right into practice Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua* (2005-2009)

<https://www.government.nl/documents/reports/2012/10/31/turning-a-right-into-practice-impact-evaluation-of-the-ixchen-centre-for-women-cervical-cancer-programme-in-nicaragua-2005-2009>

Latas C. H. (2016), *Informe de la evaluaciòn externa del Proyecto - Servicio para la inserciòn laboral de personas con discapacidad en Ecuador: Trabajando por la inclusiòn, Fase II*, (AECID, COCEMFE, FENEDIF)

Montesi F. (2015), *Dentro la scatola nera: modelli innovativi di valutazione dell'impatto per gli investimenti sociali*

www.humanfoundation.it/ita/il-nostro-lavoro/pubblicazioni

Musa e Gegbe (2013), *External Evaluation Report of Sierra Leone's Youth Reproductive Health Programme* (2007 – 2012)

<http://restlessdevelopment.org/file/external-evaluation-restless-development-sierra-leone-srh-programme-2007-2012-pdf>

OCSE/DAC (1991), *Principle of evaluation of development assistance*

<http://www.oecd.org/development/evaluation/50584880.pdf>

Office, Joint Evaluation Unit (2006), *Evaluation methods for the European Union's external assistance* (vol. 1-2-3-4)

https://ec.europa.eu/europeaid/evaluation-approach-and-methodology_en

<http://ec.europa.eu/europeaid/>

Oxfam (2016), Effectiveness Review Series, *Enhancing effectiveness through evidence-based learning*

<http://policy-practice.oxfam.org.uk/publications/how-are-effectiveness-reviews-carried-out-594353>

Patton, M.Q. (2011), *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, The Guildford Press, New York, NY.

PHINEO, (2017) *Social Impact Navigator*, Edizione Italiana

https://www.phineo.org/downloads/PHINEO_Social_Impact_Navigator_ITA.pdf

Porter M.E., Hills G., Pfitzer M., Patscheke S., Hawkins E. (2012) *Measuring Shared Value How to Unlock Value by Linking Social and Business Results*, FSG

http://www.fsg.org/Portals/0/Uploads/Documents/PDF/Measuring_Shared_Value.pdf

Pretty J. (1995) *Participatory Learning for Sustainable Agriculture in World Development* 23(8): 1247-12643

Punton, M. Welle, K. (2015) *Straws-in-the-wind, Hoops and Smoking Guns: What can Process Tracing Offer to Impact Evaluation?* CDI Practice Paper 10, Brighton: IDS

www.ids.ac.uk/publication/straws-in-the-wind-hoopsand-smoking-guns-what-can-process-tracing-offer-toimpact-evaluation

Puri A., Aladysheva A., Iversen V., Ghorpade Y., Brück T., (2015), *What methods may be used in impact evaluations of humanitarian assistance?* IZA DP No. 8755

<http://ftp.iza.org/dp8755.pdf>

Raupp M., Newman B., Revés L., Lauchande C. (2015), *Impact evaluation for the USAID Aprender A Ler Project in Mozambique*

http://pdf.usaid.gov/pdf_docs/PA00KB79.pdf

Sacks, J. 2002. *The Money Trail: Measuring your impact on the local economy using LM3*. London: New Economics Foundation

Shadish W.R., Cook T. D., Campbell D. T. (2002), *Experimental and quasi-experimental designs for generalized casual inference*. Wadsworth Cengage learning

<http://impact.cgiar.org/pdf/147.pdf>

Spitz G., Muskens R., van Ewijk E., (2013), The Dutch and development cooperation

<https://www.ncdo.nl/sites/default/files/Report%20Analysis%20The%20Dutch%20and%20Development%20Cooperation%20FINAL%202013%2003%2004.pdf>

Stern E. (2015), *A guide for Commissioners and Managers*

https://www.bond.org.uk/data/files/Impact_Evaluation_Guide_0515.pdf

Stern E., Stame N., Mayne J., Forss K., Davies R., Befani B. (2012), *Broadening the range of design and methods for impact evaluation- Report of a study commissioned by the Department for international development*, working paper 38.

The World Bank (2003), *A User's Guide to Poverty and Social Impact Analysis*

http://siteresources.worldbank.org/INTPSIA/Resources/490023-1121114603600/12685_PSIUsersGuide_Complete.pdf

UNEG (United Nations Evaluation Group), (2016), *Norms and Standards for Evaluation*

<http://www.unevaluation.org/document/detail/1914>

Unione Europea (2014), *Cooperazione internazionale e sviluppo. La lotta alla povertà in un mondo che cambia*, collana Le politiche dell'Unione Europea

USAID (2011), *Evaluation Policy*

<https://www.usaid.gov/sites/default/files/documents/2151/USAIDEvaluationPolicy.pdf>

USAID (2013), *Technical notes. Impact Evaluation*

https://www.usaid.gov/sites/default/files/documents/1870/IE_Technical_Note_2013_0903_Final.pdf

Vernon B., Copps J. (2010), *The little blue book*

<http://www.thinknpc.org/publications/the-little-blue-book/>

White, H., Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework*. New Delhi: International Initiative for Impact Evaluation.

White H., Sabarwal S., de Hoop T. (2014), Randomized Controlled Trials (RCTs), *Unicef Methodological Briefs*, Impact Evaluation No. 7

https://www.unicef-irc.org/publications/pdf/brief_7_randomized_controlled_trials_eng.pdf

WFP (2016) *Technical Note, Impact Evaluation*

<http://documents.wfp.org/stellent/groups/public/documents/reports/wfp286784.pdf>

