Independent Evaluation Office

OCCASIONAL PAPER SERIES

Empowered lives.
Resilient nations.

A Realist Review of Climate Change Adaptation Programme Evaluations – Methodological Implications and Programmatic Findings

These occasional papers from UNDP's Independent Evaluation Office (IEO) are a series of ad hoc publications presenting in edited form, papers on evaluation topics. Occasional papers cover one or more of the main themes of the Office's work, such as assessment of development results, thematic evaluations, global or regional programme evaluations, decentralized evaluations, impact evaluations, evaluations commissioned by programme units, evaluation methods and guidance, and national evaluation capacities development. They are intended primarily for international organizations, government, civil society and the academic community.

Author(s):
Takaaki Miyaguchi, Associate Professor, College of International Relations, Ritsumeikan University
Juha I. Uitto, Director, Independent Evaluation Office, Global Environment Facility

Copyright © 2015

United Nations Development Programme

1 UN Plaza, New York, NY 10017, USA

The views expressed in this publication are those of the author(s) and do not necessarily represent those of the United Nations, including UNDP, or the United Nations Member States.

All rights reserved. No part of this paper may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission. To seek permission, please email ieo@undp.org.

TABLE OF CONTENTS

ABSTRACT 1

I. INTRODUCTION 2

II. APPROACH AND METHODOLOGY 3

III. FINDINGS 9

IV. METHODOLOGICAL IMPLICATIONS AND LIMITATIONS 15

V. CONCLUSIONS 18

REFERENCES 19

ABSTRACT

Evaluating the effectiveness of climate change adaptation interventions has proven to be a difficult task. In an effort to shed light on the subject, this paper presents the methodology and results of a realist review of a set of climate change adaptation (CCA) programmes and their evaluations commissioned by the UNDP. This paper analyses CCA programmes in nine countries: Armenia, Egypt, Malawi, Mozambique, Namibia, the Philippines, Tanzania, Turkey and Zimbabwe. Together with their respective host governments, these programmes were implemented by either UNDP or various United Nations partner agencies and have already been evaluated. The authors have used a realist approach and conducted a meta-analysis of the programmes' evaluation reports. Based on the analytical frameworks for evaluating CCA interventions, as well as the authors' own field experience, the paper hypothesizes a number of key context, mechanism, and outcome configurations, which are considered vital in a realist approach, but have not yet been widely tested in the field of CCA. Although they encountered limitations and methodological challenges, the authors posit that adopting a realist approach to complex development projects, such as these CCA programmes, is indeed a useful way of providing applicable explanations, rather than generalizations or judgments, of what types of projects/activities work for whom, in what circumstances, and how, for future CCA interventions in developing countries.

I. INTRODUCTION

It is now clear that climate change is a reality. Vulnerable communities, especially in developing countries, have been increasingly exposed to climate-related hazards. Although the link between climate change and disasters remains scientifically unproven (Pielke, 2014), climatic hazards take a high toll and have the ability to set back development considerably. It has been estimated that four billion people are vulnerable and 500 million people are at extreme risk from forms of hydro-meteorological disasters and other climate risk, including floods, droughts, landslides and tropical cyclones (Global Humanitarian Forum, 2009).

However, mitigation interventions against climate change, e.g. efforts to reduce greenhouse gas emissions, have received more attention and funding from donor countries than climate change adaptation (CCA). CCA interventions aim to lessen climate-related vulnerability and build so-called 'adaptive capacity' in vulnerable communities. But CCA has not been the main focus of the donor community until recently, despite the fact that the CCA issues very much pertain to the lives and livelihoods of people and communities.

In recent years, though, more attention has been paid to the CCA agenda. Consequently we have seen the emergence of a number of CCA programmes in developing countries, which symbolizes an important shift in the donor and international community. Many of the CCA programmes that began several years ago have now been completed and evaluated.

There are several types of adaptation: autonomous and planned (Pittock & Jones, 2000); reactive and proactive/anticipatory; short-term or longer term; localized or widespread (Smit, Burton, Klein, & Wandel, 2000); and targeted and integrated (Agrawal, 2008). But ultimately, the overall objectives of CCA interventions are to lessen the impact of climate change through soft and hard interventions and to increase the adaptive capacity of vulnerable people, policymakers, organizations, technicians and others.

This paper analyses CCA programmes that have: (1) been implemented by UNDP and other United Nations agencies; (2) finished initial implementation cycles; and (3) been evaluated. The research takes a realist approach to a meta-analysis of the final evaluation reports of these CCA programmes. This paper presents both methodological implications and programmatic recommendations coming out of such evaluations.

II. APPROACH AND METHODOLOGY

Since about 2011, implementation of some of the first United Nations CCA programmes has ended. Although there are some notable studies (e.g. Hedger, Horrocks, Mitchell, Leavy, & Greeley, 2010; Bours, McGinn, & Pringle, 2014; Valencia, 2010; Frankel-Reed, Brooks, Kurukulasuriya, & Lim, 2010), the volume of literature related to evaluation of CCA programmes is still limited. Evaluation of CCA programmes is still relatively new and comes with a number of challenges. There is some discussion of whether CCA evaluation requires approaches and criteria that are different from regular development evaluation. For example, Picciotto (2007) has argued that climate change places more emphasis on risk and uncertainty. Adger, Arnell, & Tompkins (2005) list four evaluation criteria for CCA interventions: efficiency, effectiveness, equity and legitimacy. Frankel-Reed, Brooks, Kurukulasuriya, & Lim (2010) identify conceptual challenges including: (1) varying definitions and scope of adaptation; (2) fuzzy delineation between 'business-as-usual' and 'additional' adaptation interventions; and (3) the very nature of adaptive capacity and how to measure and evaluate it. These are further complicated by both geographical (coverage) and temporal (short to long-term) scales. Bours, McGinn, & Pringle (2014) summarize various monitoring and evaluation frameworks and concepts proposed by development agencies and institutions, providing a guiding map for the evaluation of CCA programmes. Given the growing importance of CCA programmes in the international development field, there is an urgent need to further promote evaluation studies of CCA. This paper aims to contribute to the field through a realist review of final evaluation reports of those CCA programmes.

REALIST APPROACH

The realist approach belongs to the school of theory-based evaluation (Stern, et al., 2012). According to Pawson & Tilley (2004), the approach emphasizes underlying assumptions about the way certain interventions are expected to yield certain outcomes. Theory-based evaluation is guided by a programme theory that intends to explain how an intervention is to achieve its outcomes. The programme theory helps evaluators and implementers to explain not just *what* the programme is expected to achieve, but also *how* it is expected to achieve it (Weiss, 1997). This approach has been long adopted by various development agencies, often in the form of a logical framework. The importance of a theory of change (TOC) has been emphasized as a key tool for evaluating environment and development interventions (Uitto, 2014; Vaessen & Todd, 2008). A TOC is to make sure that the underlying assumptions through which desired changes are triggered and realised are made explicit, often highlighting the contextual conditions that may influence the intervention's outcomes or results (Funnell & Rogers, 2011; Weiss, 1997). Accordingly, TOC is considered a useful approach in evaluating complex international development projects and interventions (Center for Global Development, 2006).

Within the broader theory-based approach, the realist approach has been put forward by Pawson & Tilley (1997), initially based on practices in criminology and health studies. Although disentangling the practical meaning and application of the realist approach poses methodological and practical challenges (e.g. Betts, 2013; Greenhalgh, Wong, Westhorp, & Pawson, 2011), the realist approach is, according to the words of Rycroft-Malone, et al. (2012):

> "philosophically rooted in realism, which combines three social science principles: causal explanations are achievable; social reality is mainly an interpretative reality of social actors; and social actors evaluate their social reality. Realism involves identifying underlying causal mechanisms and exploring how they work under what conditions. This contextually bound approach to causality is represented as context + mechanism = outcome. Therefore, it is an intuitively appealing approach to those trying to expose and unpack the complexities of contexts and interrelated mechanisms underlying implementation activity."

The realist approach is about (programme) theories that offer 'explanations' rather than judgments. Through this approach, an evaluator does not simply advise policymakers *whether* an intervention works, but can help discover why, when and how it might succeed (Pawson, Greenhalgh, Harvey, & Walshe, 2004). Moreover, this approach

does not try to examine study by study or intervention by intervention, but programme theory by programme theory (*ibid.*). Since it is not possible to review every single programme theory, a realist review should "prioritize which programme theories to be inspected" (*ibid.*).

A realist approach is believed to be more suited for evaluating complex programmes with different causal mechanisms operating in different contexts (Stern, et al., 2012). The approach is thus more suited to international development, where such development programmes are often characterized by so-called 'high causal density' (Woolcock, 2013). High causal density accompanies high uncertainty coming from the differing types and interests of stakeholders and openness to exogenous influences. Complex development programmes normally accompany interventions that cannot or should not be viewed by themselves without consideration of direct, indirect or potential influence from varying contexts. This task of identifying context resonates with the 'horizontal search' proposed by Cartwright & Hardie (2012), which is to look at the support factors that enable certain policies and mechanism to produce results.

However, adopting a realist approach is not always straightforward in practice. Although Pawson, Greenhalgh, Harvey and Walshe (2004) lay out a sequence for realist review, in effect, adopting a realist approach "employs no one standard 'formula', other than the base strategy of producing a clear theory of programme mechanisms, contexts and outcomes" (Pawson & Tilley, 1997). In essence, identifying CMO (context, mechanism, outcome) configurations is the vital element of realist approach.

CMO CONFIGURATIONS

In identifying how an intervention achieved its outcomes (O), a realist evaluator examines both the underlying mechanisms (M) and the context (C). Realist approach is "all about hypothesizing and testing CMO configurations" (Pawson & Tilley, 2004). The importance and significance of CMO configurations can be highlighted through comparison with a deterministic evaluation approach, which seeks specific underlying mechanisms that yield fixed outcomes. A deterministic approach sees such mechanisms externally valid, regardless of the differing contexts. This is in effect telling policymakers *whether* an intervention works while ignoring the importance of the context. A realist approach, on the other hand, pays utmost attention to the contextual conditions. CMO configurations are a conceptual matrix through which selected programme theories are expressed. Programmes work (i.e. have successful outcomes) only when programme activities introduce the appropriate ideas and opportunities (i.e. mechanisms) to certain groups of people in appropriate social and cultural conditions (i.e. contexts) (Pawson and Tilley, 1997).

Mechanisms in the CMO configuration concept refer to a possible line of causation generated by programmes and interventions. Mechanisms are thus not the interventions themselves (or programme activities), but the way that interventions create change. Pawson and Tilley (1997) draws an example of a clock and the way it works. It is never possible to understand how it works simply from the movement of its hands or an examination of its face (i.e. intervention activities). An understanding of the construction of the balanced spring or oscillation of caesium atoms (i.e. underlying mechanisms) is needed to understand how it works. The concept of mechanism in a realist approach is the acknowledgement of the stratified reality that surrounds a programme. It asks what it is about a programme that makes it work, as opposed to simply asking *whether* a programme works. Mechanism (M) thus consists of two components: intervention activities and programme theory. Intervention activities are represented as implementation theory, and programme theory represents the mechanism of change (Weiss, 1997). Implementation theory and programme theory together comprise the Theory of Change and thus mechanisms.

Outcome refers to the intended or unintended consequences of programmes. However, an outcome does not simply refer to a measured change of altered behaviour. Rather, it can take many forms depending on the different time scales, types of discerned mechanisms, contexts and so forth (Pawson & Tilley, 2004). The concept of outcome is used to describe the change in rates between one regularity (e.g. baseline condition before the introduction of programme activities) and another (e.g. altered status of condition, behaviour, or situation). But this does not mean that a realist evaluation looks for one fixed type of outcome pattern. It aims to decipher the reasons *why*, *how*, and *when* certain outcome patterns emerge depending on the context, instead of focusing on and measuring fixed outcome patterns.

Context refers to the external and internal circumstances that render (or not) those mechanisms active. A realist notion states that the relationship between mechanisms and outcomes is not fixed, but contingent (Pawson and Tilley, 1997). Certain mechanisms generate certain outcome patterns, but only in certain contexts. Cartwright and Hardie (2012) draw upon an actual development project, the Bangladesh Integrated Nutrition Project (BINP), to assert the importance of contextual conditions. The design of BINP was heavily based on an earlier project in neighbouring India, the Tamil Nadu Integrated Nutrition Project, which was deemed successful. BINP was to provide nutritional counselling for pregnant women, and supplementary feeding for children under 24 months who were especially deprived. Ultimately, though, BINP was unsuccessful. In contrast to the Tamil Nadu case, BINP used food not as a supplement, but as substitute (so that other members of the family could be fed). Also, it was often the case that men or mothers-in-law were the decision-makers instead of mothers, which further diluted the efficacy of project interventions (Cartwright and Hardie, 2012). Discovering such pertinent contextual conditions is vital in proving the external validity of programme interventions and is the key in a realist approach.

A realist approach, then, does not take a static or deterministic view of whether or not a programme works (i.e. paying attention only to the M-O relationship). It seeks to understand how (i.e. underlying mechanisms) the programme works and under what conditions (i.e. contexts). In summary, a realist review encompasses "theory-driven interpretative techniques, which were developed to help make sense of heterogeneous evidence about complex interventions applied in diverse contexts in a way that informs policy" (Greenhalgh, Wong, Westhorp, & Pawson, 2011).

STUDY MATERIAL

The material used for analysis in this study consisted of the terminal – or final – evaluation reports of CCA programmes and projects implemented by UNDP and its partners in the United Nations system. In 2013, the first batch of 11 CCA programme and project evaluation reports was submitted to UNDP's Independent Evaluation Office (IEO). What is unique about these is that they are the first evaluation reports of completed CCA programmes within the UNDP system.

IEO receives and assesses the quality of evaluation reports prepared and submitted by country offices and other programme units (so-called 'decentralized evaluations'). The quality assessment is conducted with reference to the key assessment parameters: (1) terms of reference; (2) evaluation subject, context and purpose; (3) evaluation framework; (4) findings; (5) conclusions; and (6) recommendations and lessons. The rating scales for these parameters are: Highly Satisfactory; Satisfactory; Moderately Satisfactory; Moderately Unsatisfactory; Unsatisfactory; and Highly Unsatisfactory. Of the 11 CCA evaluation reports, nine were assessed by IEO to be moderately satisfactory or higher. Being rated moderately satisfactory or better, however, does not mean that the programme (interventions) itself was successful. Rather, it means that the evaluation report was considered to be of high quality. So high quality evaluation reports can equally cover unsuccessful programme results and achievements. For this realist review, we have decided to include only those evaluations whose quality was rated moderately satisfactory or better. It was to ensure that the evaluations that meet the specific set of standards to allow equal comparison were analysed.

These nine programmes were implemented in nine countries: Armenia, Egypt, Malawi, Mozambique, Namibia, the Philippines, Tanzania, Turkey and Zimbabwe. They represented a diverse group in terms of funding sources and budgets (such as Global Environment Facility, Millennium Development Goals Achievement Fund, and United Nations internal resources); types of beneficiaries, target audiences and regions (ranging from local vulnerable communities to inter-ministerial mainstreaming at the government level); and implementation modalities (including UNDP stand-alone, United Nations interagency joint programming and Delivering as One) (Table 1).

Since all of these evaluations were commissioned by different entities (mostly UNDP country offices) and written independently by consultants at the country level, the terms of reference, presentation styles and contents varied considerably. However, there are four common sections/criteria that all the reports covered: (1) relevance; (2)

efficiency; (3) effectiveness; and (4) sustainability. Our analysis considered these four evaluation criteria separately in order to answer these key questions:

- What are the important 'regularities' (outcomes) recognized by the evaluators of the CCA programmes for each evaluation criterion?
- What are the underlying mechanisms that increase or decrease those regularities?
- What are the contextual conditions that necessarily enable or foster the mechanisms to generate desired outcomes?

Table 1: List of the CCA programme/project evaluation reports reviewed

Country	Programme/Project Title	Duration (months)	Implementation Modality (Funding Source)
Armenia	Adaptation to Climate Change Impacts in Mountain Forest Ecosystems of Armenia	May 2009 - Jun 2013 (50m)	UNDP (GEF)
Egypt	Joint Programme: Climate Change Risk Management in Egypt	Oct 2008 - Apr 2013 (55m)	JP (MDG-F)
Malawi	The National Programme for Managing Climate Change in Malawi	Apr 2010 - Dec 2012 (33m)	UNDP (AAP)
Mozambique	Joint Programme on Environmental Mainstreaming and Adaptation to Climate Change in Mozambique	Sep 2008 - Aug 2012 (48m)	JP (MDG-F)
Namibia	Namibia Country Pilot Partnership Programme; Adapting to Climate Change through the Improvement of Traditional Crops & Livestock Farming	Jun 2007 - Dec 2011 (55m)	UNDP (GEF)
Philippines	Joint Programme: Strengthening the Philippines' Institutional Capacity to Adapt to Climate Change	Dec 2008 - Dec 2011 (37m)	JP (MDG-F)
Tanzania	Joint Programme on Environment with a Focus on Climate Change, Land Degradation/Desertification and Natural Resources Management	Oct/Dec 2009 - Jun 2011 (21m)	JP (MDG-F)
Turkey	Joint Programme on Enhancing the Capacity of Turkey to Adapt to Climate Change	Apr 2008 - Dec 2011 (45m)	JP (MDG-F)
Zimbabwe	Coping with Drought and Climate Change in Zimbabwe Project	May 2008 - Sep 2012 (53m)	UNDP (GEF)

JP – Joint Programme; **MDG-F** – Millennium Development Goal Achievement Fund; **GEF** – Global Environment Facility; **AAP** – Africa Adaptation Programme

ANALYSIS METHOD

All evaluation reports covered the common evaluation criteria of relevance, efficiency, effectiveness, and sustainability. The definitions of the four criteria are based on the OECD's Glossary of Key Terms in Evaluation and Results-based Management (2002) adapted by UNDP and its partners (Table 2).

Table 2: Definitions of Key Evaluation Criteria

Criteria	OECD Definition
Relevance	The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies.
Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
Effectiveness	The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
Sustainability	The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.

Source: OECD (2002)

The authors of this paper sought to unearth, hypothesize and construct CMO configurations, in accordance with the procedures of a realist review. Because all nine evaluation reports discuss the four common criteria separately, separate CMO configurations were sought for each.

To help identify CMO configurations, the authors recorded every evaluative remark found in the reports and assigned either a positive (+) or negative (-) attribute. This step generated a total of 577 items. Each item thus belongs to one of the four evaluation criteria. The itemised remarks within each evaluation criterion were then clustered by several key components. Key components are those considered to play an important role in achieving a high (or low) level of relevance, efficiency, effectiveness or sustainability. These key components were gleaned from the reports. Table 3 shows some of the key components identified through this process.

Table 3: Some key components for each evaluation criterion

Criteria	Identified Key Components
Relevance	Community/district, government, donor/global level alignment and relevance
Efficiency	Financial execution, programme management, staff hiring and retention, stakeholder involvement
Effectiveness	Adaptive capacity, adaptive measures, mainstreaming, awareness raising
Sustainability	Long term adaptive capacity, adaptive measures, initiatives and replications, long-term stakeholder engagement, mainstreaming

An important distinction that needs to be made clear is the treatment of CMOs' O (outcome). Outcomes within CMO configurations that Pawson and Tilley (1997; 2004) refer to are those that together contribute to the achievement of the programme's ultimate goal. In this review, Os refer to the key components that together contribute to high/low levels of relevance, efficiency, effectiveness or sustainability, and not to the ultimate CCA programme goals. Not treating Os as the outcomes that lead to the programme goal is one limitation that this realist review has faced (more on limitations in a later section).

The M (mechanism) of CMO configurations represents theory of change, which comprises both implementation theory (intervention activities) and programme theory (the mechanisms of change). The authors aimed to identify key programme theories by examining the relationship between intervention activities and key components' regularities. Determining M is to answer what it is that allows certain programme activities to influence the achievement of the four evaluation criteria. Contexts (C) were identified through the study of the various elements that surrounded the CCA programmes, e.g. target populations, stakeholders, government implementing partners, funding schemes, macroeconomic and socioeconomic situations.

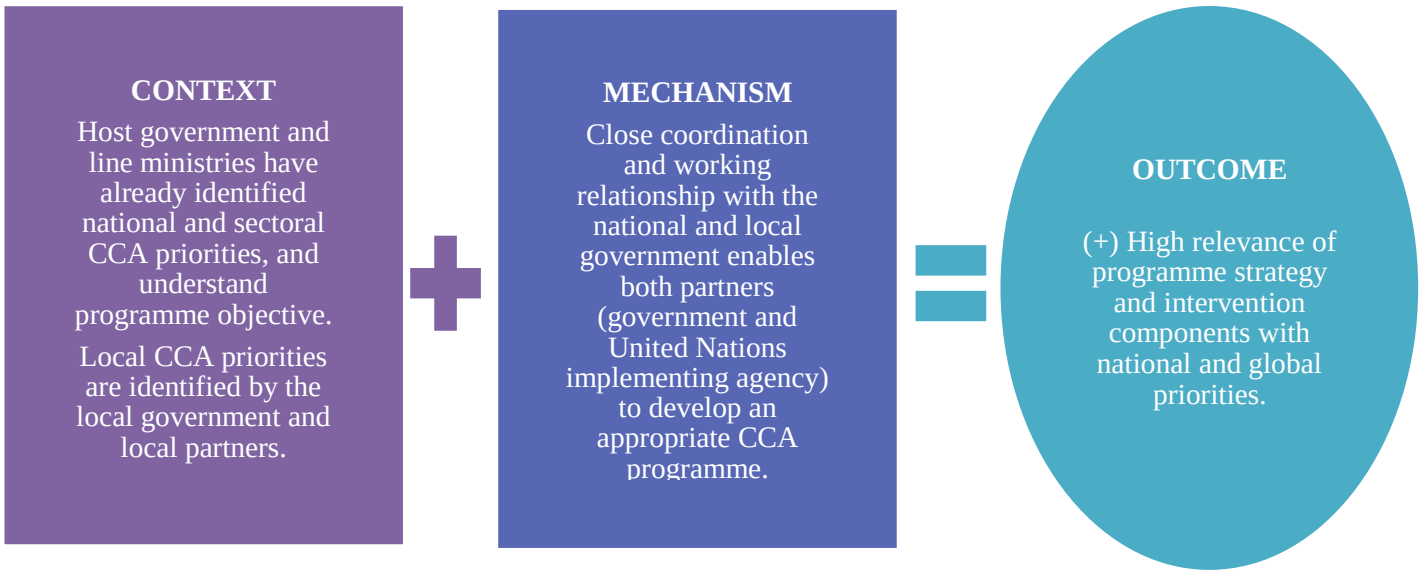
III. FINDINGS

RELEVANCE

For relevance, identifying potential M-O configurations was straightforward. That is because this criterion relates simply to the level of alignment of the CCA programme design and activities with government and global priorities. A positive outcome (O) in this criterion is therefore a high level of alignment. A mechanism (M) that can help generate such outcome patterns can be close coordination with the host government and the line ministries, because this enables both partners (government and United Nations implementing agencies) to develop an appropriate CCA programme. Table 4 lists CMO configurations for the relevance criterion.

For the implementing United Nations agencies to develop a CCA programme that is consistent with the country's national, local and global priorities, there must be close coordination and working relationships with the central and local governments in the target country. However, depending on the contextual conditions, this otherwise apparent underlying mechanism does not always generate a highly relevant programme design. There were cases where such a mechanism did not seem to have worked. In one case, CCA priorities in a certain region were not identified before or during the programme inception stage (Mozambique). In another case, the government counterpart did not fully understand the objectives of the CCA programme although national and local priorities had been identified (Turkey). Under these different contextual conditions, programme theory (or mechanisms of change) did not seem to work, even though implementation theory (intervention activities) was well placed.

Table 4: CMO Configuration for Relevance Criterion



In evaluation, relevance is often monitored throughout a project's lifetime to see whether or not the project objectives are still aligned with national and global priorities amid changing circumstances. But relevance can also represent the relationship between outcomes and impact (van den Berg, 2011). Since these evaluation reports were written immediately before or after implementation closure, the evaluators could only estimate the likelihood of programmes moving towards impact, partially by drawing on the level of relevance. A less than high level of relevance can negatively influence the likelihood of achieving impact in the future. The relevance criterion should thus be viewed as an important pathway towards future impact.

EFFICIENCY

For the efficiency criterion, multiple types of outcomes (or key components that can further lead to high/low degree of efficiency) were identified. The efficiency criterion looks at the relationship between inputs and outputs. Table 5 shows the selected, key CMO configurations for efficiency criterion.

First outcome patterns included high staff turnover rate, which severely affected the level of efficiency of programme implementation. Two unintended consequences, namely a delayed recruitment process and relatively short programme duration, contributed to lowering the morale of hired staff and their career expectations. Delayed recruitment process was reported in three countries (Zimbabwe, the Philippines and Malawi). Relatively short programme duration was reported in Tanzania (with 21 months of programme duration), followed by Malawi (33 months) and the Philippines (37 months). These short-lived programmes were also those that suffered from high rates of staff turnover. However, the specific contextual conditions that might have exacerbated this negative mechanism of change may include a high level of demand on jobs in international development within the country as well as the domestic market. Conversely, considering that Zimbabwe has been experiencing economic turmoil, and consequently has limited job opportunities in international development as well as in the domestic market, it seems to make sense that Zimbabwe had a lower rate of staff turnover than other countries with similarly short programme durations and delayed recruitment. It can then be hypothesized that, under such circumstances, people who could secure a position from a donor-funded project tended to want to remain in the position no matter what might happen to the programme activities. Such hypothesis, however, should be treated carefully in this type of realist review, since it inherently contains the risk of overlooking important macro- as well as micro-level circumstances that may otherwise influence outcome patterns. This is one of the limitations that we have faced in this review, which is discussed further in a subsequent section.

The second key outcome patterns relate to a high level of stakeholder involvement. Building trust and good working relationships with relevant stakeholders at an early stage seems a sure way of increasing the motivation level amongst stakeholders who are thus more likely to participate in the programme. The programme theory here is that building trust and good working relationships will increase the level of motivation among active participants, which leads to high stakeholder involvement. The contextual conditions necessary for this M-O configuration to be effective may be a focused programme design, such as mountain ecosystem protection in Armenia, traditional crops and livestock farming in Namibia, and targeting rural farmers in one district of Zimbabwe. But a narrow programme scope does not suggest an automatic boost in stakeholder involvement. Without the necessary programme activities and mechanisms of change, different outcome patterns may be generated.

Table 5: Key CMO configurations for efficiency

CONTEXT		MECHANISM		OUTCOME
A high level of demand on jobs in international development and the domestic market	+	Delayed recruitment process and relatively short programme duration make hired staff lose their work and career motivation	=	(-) High staff turnover rates
Sector specific and focused programme design Relevant stakeholders are supportive of United Nations and well aware of CCA issues and risks	+	Partnerships with stakeholders are built at an early stage, where they feel more motivated to participate in the programme	=	(+) High stakeholder involvement
Strong leadership from national executing agency Sector and region specific scope of programme	+	Adaptive management and clearly defined roles and responsibilities to each party enable the programme to attend to the needs and demands of the local beneficiaries whilst maintaining the ultimate programme goal	=	(+) High level of programme management achievements
Government and implementing United Nations agencies have not worked or reached an agreement on a common financial disbursement system	+	Working with multiple United Nations agencies for joint implementation with tight/short programme schedules makes it difficult to ensure financial accountability and sound financial management	=	(-) Low level (both quality and quantity) of financial execution

The third outcome patterns identified were a high level of programme management achievement, such as swift financial execution and implementation of activities. A mechanism identified in generating this outcome pattern incorporates implementation of adaptive management with clearly defined roles and responsibilities for each involved party. That, in turn, can create an enabling environment for the programme to respond to the needs and demands of the local beneficiaries. This positive M-O relationship seems to have been facilitated by strong leadership of the national executing agency and government counterparts in charge of implementation.

The fourth outcome patterns are concerned with a low level (quality and quantity) of financial execution. It may seem obvious that having to work with multiple United Nations agencies for joint implementation with tight schedules makes it more difficult to ensure sound financial management. However, such a mechanism is not a deterministic contextual condition for generating a negative outcome. This negative outcome can be avoided by a mutual agreement between government and implementing United Nations agencies on a common financial disbursement system (Armenia and Namibia).

EFFECTIVENESS

Evaluation of the effectiveness criterion can be the most important indicator for a 'successful' CCA programme, because it determines some of the crucial outcome patterns, such as adaptive capacity (the capacity of the relevant stakeholders to cope with and reduce the risk posed from climate change), adaptive measures (from policy formulation to new farming and climate modelling techniques), and mainstreaming (among ministries, sectors and stakeholders). Table 6 shows the key CMO configurations identified for the effectiveness criterion.

The evaluators had differing views on what constituted 'high effectiveness'. Some evaluators recognized mere delivery of simple outputs as evidence of high effectiveness, while others looked beyond simple output delivery levels. Nonetheless, three important key CMO configurations were constructed for this criterion.

Table 6: Selected CMO configurations for the Effectiveness criterion

CONTEXT		MECHANISM		OUTCOME
Specifically identified types of participants are well aware of the climate risks Specific types of skills that they need to acquire are clear to them	+	Training and transfer of needed techniques and practices for the relevant people facilitate these skills, techniques and knowledge to be applied and used	=	(+) High level of adaptive capacity (+) High level of utilisation of adaptive measures
Relevant ministries and stakeholders are highly aware of the climate risks and the vital importance of reducing vulnerability	+	Provision of relevant technical, policy and advisory support to relevant people (from government staff to rural farmers) facilitates its integration with their "business-as-usual" activities	=	(+) Wide range of mainstreaming
General citizens are relatively unaware or lack knowledge of climate change and associated risks Government officials do not see the actual need to integrate CCA issues in their business-as-usual activities	+	TV, newspaper and symposium for wider publicity attract attention and boost curiosity in citizens about CCA issues Dissemination targeting government officials, even with convincing technical data and studies, does not itself alter the behaviour of government officials	=	(+) Raised level of awareness amongst the general public (-) Low level of awareness amongst government officials

First outcome patterns are both a high level of adaptive capacity and utilization of adaptive measures. Training and transfer of required techniques and practices to the relevant people are the intervention activities themselves. The mechanism of change (or programme theory) would be that such acquired skills, techniques and knowledge are applied and used by the end-users, be they farmers or government officials. Contextual factors, however, are important for this M-O relationship to work (for example, the specificities of the beneficiary groups; their awareness of climate risks; and the types of skills, techniques and technologies that are needed and desired). Without such contexts, the mechanism does spark the generation of positive outcome patterns.

The second outcome patterns include a wide range of mainstreaming efforts. The programme gave relevant technical, policy and advisory assistance to relevant people (be they government officials or rural people), thereby

generating enthusiasm for integrating and utilizing such assistance in their business-as-usual work or activities. But the mere provision of assistance did not automatically trigger integration or utilization by end-users. For this mechanism to be effective, relevant ministries and stakeholders should be highly aware of climate risks and of the vital importance of reducing vulnerability.

The third outcome patterns are closely related to the second outcome patterns. This outcome is about a high level of awareness in the general public. Under circumstances where the general public is known to lack awareness or knowledge of climate change and associated risks, means of dissemination such as television, social media, internet or newspapers may be able to boost their curiosity and awareness. However, when it comes to raising awareness among government officials, this seemingly simple mechanism of dissemination has not been found to alter behaviour or mind-sets unless they see and understand the actual need to integrate CCA issues with their business-as-usual activities at the government level.

SUSTAINABILITY

The fourth and last evaluation criterion is sustainability. In extracting key CMO configurations for this criterion, it has proved to be quite challenging to differentiate between effectiveness and sustainability. This is partially due the fact that the CCA evaluation reports were written immediately before or after the programme completion. That makes it difficult to properly evaluate sustainability, which can only be verified over the longer term. One can thus say that the related CMO configurations are the most context-dependent. Nonetheless, the key CMO configurations are listed in the table below (Table 7).

First, outcome patterns that are vital to achieving sustainability are high likelihood for sustaining adaptive capacity and a sustained high utilization level of adaptive measures introduced by the programme. To achieve this outcome, adaptive capacity and measures should be those demanded by the end-users/beneficiaries themselves. For example, both weather forecasting technology and related capacity development (Egypt) and new farming practices and techniques (Zimbabwe) are among the adaptive capacity and adaptive measures clearly demanded by end-users. Other programmes did not know or simply assumed a demand for the types of training and other measures they were introducing. The latter situation may not foster a sense of ownership towards the adaptive capacity and measures introduced. A subtle, but important, difference is placed between ‘necessary’ and ‘requested’ types of adaptive capacity and measures. The former, though theoretically and technically correct, may not foster a sense of ownership, while the latter is an important prerequisite for generating positive sustainability outcomes. Important contextual conditions may therefore include that: (1) key government counterparts, end-users and beneficiaries are already highly aware of the CCA programme's intended objectives; (2) they have clear ideas as to what types of adaptive capacity or measures they need; and (3) there is sustained political interest in the CCA programme's intended objectives.

Secondly, key outcome patterns are concerned with a high likelihood of sustained stakeholder engagement. A high level of stakeholder engagement can be triggered through formulation of communities of practice, such as associations, networks, partnerships or committees among key stakeholders. Such communities of practice provide a useful platform for committed partners/stakeholders to continue to be active. However, for this M-O configuration to hold requires government officials and beneficiaries on the ground to be active and continue to see the need for, and benefits from, engaging with the CCA programme and its ultimate objectives.

The third set of outcome patterns are concerned with a sustained high level of mainstreaming at central policy and planning level. This is applicable especially for those CCA programmes that aim to help raise adaptive capacity at a national level. The relevant programme activities included, for example, establishment of a meteorological unit in Zimbabwe and the development of a cross-sectoral plan and policy framework in Tanzania. When these activities were implemented within the existing local/national and institutional frameworks, they seem to have harnessed a sense of ownership and thus triggered smooth integration into the planning and policies of the respective institutions. But that mechanism is only valid when government counterparts understand the need for mainstreaming. Other important contextual conditions identified are sufficient institutional resources and motivation of government officials.

Table 7: Selected CMO configurations for Sustainability

CONTEXT		MECHANISM		OUTCOME
Key government counterparts, end-users and beneficiaries have relatively high levels of understanding of CCA programme's intended objectives, and have clear ideas as to what types of adaptive capacity or measures they need Sustained political interest towards the CCA programme's intended objectives	+	Development of adaptive capacities and introduction of new adaptive measures that are requested by the end-users and can yield tangible results foster a sense of ownership towards built capacities and introduced measures	=	(+) High likelihood for sustaining built adaptive capacity (+) High likelihood for sustained high utilisation level of adaptive measures introduced
Beneficiaries on the ground and government continue to be present and see the need and benefits in engaging themselves to the CCA programme's intended objectives	+	Formulation of communities of practice for developing and implementing new initiatives provides a useful platform for the committed partners/stakeholders to continue to be active for the CCA matters	=	(+) High likelihood for sustained, high level stakeholder engagement
Government counterparts understand the need of mainstreaming Institution's sufficient resources and motivation level of government officials	+	Programme activities implemented within the local/national and institutional existing framework foster a sense of ownership and trigger smooth integration of planning and policies	=	(+) High likelihood for sustained level of mainstreaming at central policy and planning level
Relevant stakeholders have strong sense of ownership and have adequate resources and capabilities	+	Introduction of adaptive measures to the stakeholders and institutions with relevant mandate enables 'rooting' of these measures inside the respective stakeholders and institutions	=	(+) High likelihood of generating self-initiatives and replications in the long term

The fourth set of outcome patterns are about the likelihood of replication. For this outcome to happen, introduction of adaptive measures to the stakeholders or institutions that have a CCA-relevant function or mandate needs to take place. Such an introduction helps realise its 'rooting' and integration inside their institutions. Related examples include an additional climate change research function being integrated in an existing research centre (Egypt); a new certification course introduced by a university (Turkey); and easy-to-use farming techniques for rural farmers (Mozambique). Relevant stakeholders should have a strong sense of ownership and have adequate resources and capabilities for the mechanism to generate the outcome.

IV. METHODOLOGICAL IMPLICATIONS AND LIMITATIONS

METHODOLOGICAL IMPLICATIONS

Presented above are the results of a review of the CCA evaluation reports using a realist approach and principles. There are important methodological implications emanating from this exercise. First, this review has systematically emphasized identification of contextual conditions as part of the efforts to hypothesize CMO configurations. The design and implementation of CCA programmes are by nature quite complex. They involve a number of stakeholders, implementing partners, types of beneficiary, funding sources and requirements, and differing programme goals and local priorities. For such complex programmes, unearthing the contextual conditions is particularly vital to increase the level of external validity. Although it is clear that there will not be a panacea or one-size-fits-all CCA programme in the future, paying close attention to contextual conditions does offer a useful path for looking at some of the important elements for better CCA programming.

Second, had this review taken a more judgmental approach, it would have focused only on judging if certain CCA programme activities worked without examining the contextual conditions. With the presented CMO configurations in hand, one can systematically try to answer such questions as 'for whom does the programme or its activities work?', 'under what circumstances?', and 'how?' These considerations are vital for policymakers, programme designers and implementers, and donors. The collection of contextual conditions according to the programme activity types can provide useful explanations of what kind of contextual conditions there should be for certain programme activities to generate positive outcome patterns. This systematic pursuit of identifying the contextual conditions is one of the advantages of the realist approach.

Third, this realist review has also proved useful in identifying key programme theories. The realist approach enables the evaluator to look for common underlying mechanisms that can cut across programme activities. This exercise is not about listing or analysing each programme activity. Rather, it aims to look at common and key programme theories. If the review tries to analyse each programme activity, it will not be able to achieve external validity.

Fourth, as a mandatory requirement, all donor-funded programmes must be evaluated against their achievements at the end of implementation. The outputs of this requirement are evaluation reports, such as those analysed here. The results of this review suggest that conducting a similar realist meta-analysis of other CCA programme evaluation reports can be useful in improving the quality and direction of future CCA programming.

The fifth important methodological implication of this review is that it did not just focus on what was successful. By systematically studying both positive and negative outcome patterns (and the mechanisms and contexts that trigger them), the realist approach has proved itself useful in facilitating learning from mistakes and failures. The contextual conditions for both successful and unsuccessful programme activities can be studied for the improvement of future CCA programming.

Sixth, seeking external validity of successful programme interventions is one of the most important challenges for evidence-based policy-making, particularly in international development. Determining how success in one place can be generalized to other locations and conditions is a sensitive process. For example, evaluation of individual small projects may not increase the degree of external validity or *generalizability*, because it may only look at specific local contexts. On the other hand, evaluation of macro-level programmes may overlook the importance of local contextual conditions. A realist review can be useful in finding externally valid programme interventions and designs for CCA by identifying the key contextual conditions that help a programme achieve its goals.

To further utilize the findings of this realist meta-analysis, we created summary tables that list the types of contextual condition according to the types of mechanism (i.e. programme activities and programme theories).

With such summarized findings, one can provide explanations of what may work, how it will work and under what circumstances.

CHALLENGES AND LIMITATIONS

This review exercise encountered challenges and limitations. These were in part related to the requirements and conditions of realist principles; but also to the study material and the way it was analysed.

First, following the principles and meeting the conditions of a realist approach was difficult. In the case of a realist synthesis, according to Betts (2013), there have been only "30 syntheses that satisfied realist synthesis criteria" according to the standards proposed by Wong, Greenhalgh, Westhorp, Buckingham and Pawson (2013), called Realist and Meta-review Evidence Synthesis: Evolving Standards (RAMESES). The realist approach has been tested mostly in criminology and health studies, but not in of international development – let alone for a CCA. This review of CCA evaluation reports is one of the few that has sought application of realist principles to highly complex international development programmes. It may therefore have failed to satisfy some important aspects of realist principles.

One such aspect is related to the treatment of outcomes. Because of the way the CCA evaluation reports were written, i.e. according to the pre-set evaluation criteria of relevance, efficiency, effectiveness and sustainability, this review could only seek CMO configurations *within each criterion*. It means that outcomes used in the CMO configurations of this review do not translate into these outcomes contributing to the achievement of overall CCA programme goals (e.g. adaptive capacity built, CCA policies mainstreamed). Instead, outcomes in this review represent the key components in achieving a high/low level in the four evaluation criteria. For the efficiency criterion, for example, the authors attempted to discover the important outcomes (the key components) that may enhance the level of efficiency, under what circumstances and through what type of underlying mechanisms.

Second, an alternative method of analysis might have looked at the nine CCA programmes according to the different *programme outcomes* for achieving programme goals (as opposed to treating outcomes as key components in achieving a high/low relevance, efficiency, effectiveness and sustainability). Such an alternative analysis however would have had to incorporate several evaluation categories of CCA programming, such as:

(1) five adaptation programming components, i.e. policymaking and planning; capacity development and awareness raising; information management; design and decision-making for investments; and risk reduction practices/livelihood activities and/or resource management (Frankel-Reed, Brooks, Kurukulasuriya, & Lim, 2010);
(2) six thematic areas, i.e. agriculture/food security; water resources and quality; public health; disaster risk management; coastal zone development; and natural resources management (Lim & Spanger-Siegfried, 2004);
(3) four types of evaluation indicators, i.e. coverage; impact; sustainability; and replicability (Frankel-Reed, Brooks, Kurukulasuriya, & Lim, 2010).

This realist review could not bring any of these evaluation categories into the analysis. For example, we could not look at the information management component *across* all nine CCA programme evaluation reports. Similarly, it was not possible to construct CMO configurations for each thematic area listed above. This difficulty was mainly due to how the CCA evaluation reports were written. For future research, constructing and refining CMO configurations of different CCA programmes according to such adaptation programming components may be beneficial.

Third, as some researchers have done (most notably by Rycroft-Malone, et al. 2012), a future review should conduct a thorough search of primary studies that are relevant to the key programme theories identified. This is to consider how robust the CCA programme theories were, compared with cases identified by literature review. Although it is true that there are not yet many primary studies of this sort in the field of international development (White & Waddington, 2012), it would be a meaningful exercise to conduct in the future to strengthen the argument.

Finally, another limitation that this review faced concerned the study material and how it was chosen. The CCA evaluation reports analysed were those rated moderately satisfactory or above according to the quality assessment parameters set by IEO. Although these reports were among the first evaluations of the CCA programmes implemented by United Nations agencies, it does not mean that no other CCA evaluations exist (e.g. CCA programmes funded by a national government, bilateral development agencies, or multilateral development banks). With such a limited sample size and potential bias in mind, it should be made clear that this realist review exercise is but one of many future trials to systematically identify key CMO configurations from CCA programmes.

V. CONCLUSIONS

This paper has presented a review of nine evaluation reports of those CCA programmes that were implemented sometime between 2008 and 2012. The review was based on the final evaluation reports for the programmes in Armenia, Egypt, Malawi, Mozambique, Namibia, the Philippines, Tanzania, Turkey and Zimbabwe, using a realist approach. After itemizing and analysing a total of 577 evaluation remarks found within the reports, CMO configurations were developed for each of the four evaluation criteria covered in the reports: relevance, efficiency, effectiveness and sustainability. Through this analysis, the authors identified key programme activities and subsequent mechanisms of change in generating the outcomes under certain contexts. Notably, the contextual findings have proved useful in thinking about external validity of 'successful' interventions in one place or situation to other settings.

This realist analysis is certainly not flawless and it proved impossible to adhere to all realist principles. However, it is hoped that this paper has presented a concrete example of how to conduct a realist review using existing evaluation reports. As CCA programme design and implementation are highly complex and require focused attention to the contextual conditions, by systematically exploring such contextual conditions as part of hypothesizing CMO configurations, realist review can be a useful method for providing explanations and lessons. More broadly, it is hoped that this review exercise can be of some use for further application of the theory-based realist approach in international development.

REFERENCES

Adger, N. W., N. W. Arnell, and E. L. Tompkins. 'Successful adaptation to climate change across scales.' Global Environmental Change 15, 2005: 77-86.

Agrawal, A. 'The role of local institutions in adaptation to climate change. Presented at the Social Dimensions of Climate Change Conference, World Bank March 2008.' 2008.

Betts, J. 'Aid Effectiveness and Governance Reforms: Applying realist principles to a complex synthesis across varied cases.' Evaluation 19, 2013: 249-268.

Bours, D., C. McGinn, and P. Pringle. 'Monitoring & evaluation for climate change adaptation and resilience: A synthesis of tools, frameworks and approaches.' 2014.

Cartwright, N., and J. Hardie. 'Evidence-Based Policy: A Practical Guide to Doing It Better'. Oxford University Press, 2012.

Center for Global Development. 'When Will We Ever Learn? Improving Lives through Impact Evaluation'. Report of the Evaluation Gap Working Group. Washington D.C., USA: Center for Global Development, 2006.

Frankel-Reed, J., N. Brooks, P. Kurukulasuriya, and B. Lim. 'A Framework for Evaluating Adaptation to Climate Change: Evaluating Climate Change and Development'. In Evaluating Climate Change and Development (World Bank Series on Development), edited by O. N. Feinstein, & R. D. van den Berg, 285-298. Transaction Publishers, 2010.

Funnell, S. C., and P. J. Rogers. 'Purposeful Program Theory: Effective Use of Theories of Change and Logic Models'. John Wiley & Sons, 2011.

Global Humanitarian Forum. 'The Anatomy of A Silent Crisis'. Geneva: Global Humanitarian Forum, 2009.

Greenhalgh, T., G. Wong, G. Westhorp, and R. Pawson. 'Protocol - realist and meta-narrative evidence synthesis: Evolving Standards (RAMESES)'. BMC Medical Research Methodology 11, 2011.

Hedger, M. M., L. Horrocks, T. Mitchell, J. Leavy, and M. Greeley. 'Evaluating of Adaptation to Climate Change from a Development Perspective'. In Evaluating Climate Change and Development (World Bank Series on Development), edited by R. D. van den Berg, & O. N. Feinstein, 241-261. Transaction Publishers, 2010.

Lim, B., and E. Spanger-Siegfried. 'Adaptation Policy Frameworks for Climate Change: Developing Strategies, Policies and Measures'. Cambridge University Press, 2004.

OECD. 'Glossary of Key Terms in Evaluation and Results Based Management'. OECD, 2002.

Pawson, R, T Greenhalgh, G Harvey, and K Walshe. 'Realist synthesis: an introduction. RMP Method Paper 2/2004'. Manchester: ESRC Research Methods Programme, 2004.

Pawson, R., and N. Tilley. 'Realist Evaluation'. 2004.

—. 'Realistic Evaluation'. SAGE Publications Ltd, 1997.

Picciotto, R. 'The new environment for development evaluation.' American Journal of Evaluation 28(4) (2007): 509–521.

Pielke, Jr., R. ‘The Rightful Place of Science: Disasters & Climate Change.’ Consortium for Science, Policy & Outcomes, Arizona State University, 2014.

Pittock, A. B., and R. N. Jones. 'Adapting to what and why?' Environmental Monitoring and Assessment, 2000: 9-35.

Rycroft-Malone, J., et al. 'Realist synthesis: illustrating the method for implementation research.' Implementation Science 7:33, 2012.

Smit, B., I. Burton, T. R. Klein, and J. Wandel. 'An anatomy of adaptation to climate change and variability.' Climate change 45(1), 2000: 223-251.

Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, and B. Befani. 'Broadening the range of design and methods for impact evaluations. Report of a study commissioned by DFID.' 2012.

Uitto, J. I. 'Evaluating environment and development: Lessons from international cooperation.' Evaluation 20, 2014: 44-57.

Vaessen, J., and D. Todd. 'Methodological challenges of evaluating the impact of the Global Environment Facility’s biodiversity program.' Evaluation and Program Planning 31, 2008: 231-240.

Valencia, I. D. 'Lessons on M&E from GEF Climate Change Adaptation Projects'. In Evaluating Climate Change and Development (World Bank Series on Development), edited by R. D. van den Berg, & O. N. Feinstein, 265-283. Transaction Publisher, 2010.

Van den Berg, R. 'Evaluation in the context of global public goods.' Evaluation 17(4) (2011): 405-415.

Weiss, C. H. 'Evaluation'. Pearson Publishing, 1997.

White, H., and H. Waddington. 'Why do we care about evidence synthesis? An introduction to the special issue on systematic reviews'. Journal of Development Effectiveness 4(3), 2012: 351-358.

Wong G., T, Greenhalgh, G. Westhorp, J. Buckingham, and R. Pawson. ‘RAMESES publication standards: realist synthesis’. BMC Medicine Vol11, 2013. URL: http://www.biomedcentral.com/1741-7015/11/21

Woolcock, M. 'Using case studies to explore the external validity of 'complex' development interventions.' Evaluation 19(3), 2013: 229-248.

Empowered lives.
Resilient nations.

Independent Evaluation Office
220 East 42nd Street
20th Floor
New York, NY 10017, USA

Tel. +1(646) 781 4200
Fax. +1(646) 781 4213
www.undp.org/evaluation

Connect with us: