Multilateral Organisation Performance Assessment Network

MOPAN 3.0

2017-18 ASSESSMENT CYCLE

METHODOLOGY MANUAL

TABLE OF CONTENTS

Content

Boxes

Figures

Tables

Annexes

ACRONYMS

CO	Country Office
DER	Development Effectiveness Report
DP	Direct Partner
GDI	Gender-related Development Index
GNI	Gross National Income
HDI	Human Development Index
HQ	Headquarters
HR	Human Resources
IFI	International Financial Institution
KPI	Key Performance Indicator
MI	Micro Indicator
MO	Multilateral Organisation
MOPAN	Multilateral Organisation Performance Assessment Network
OECD-DAC	Organisation for Economic Cooperation and Development – Development Assistance Committee
RBM	Results-based management
ToR	Terms of Reference
UN	United Nations
UNEG	United Nations Evaluation Group

SECTION 1:
INTRODUCTION AND OVERVIEW

Under MOPAN's reshaped assessment approach, *MOPAN 3.0*, which was first implemented in 2015-16, the Network is assessing more organisations, collecting data from more partner countries, and widening the scope of organisations assessed. MOPAN is also sharpening its focus on results and development effectiveness, while continuing to assess organisational effectiveness. Details on the specific assessment cycles can be found in Annex E and Annex F.

The core hypothesis of MOPAN 3.0 is that *if a multilateral organisation has effective systems, practices and behaviours in place, then its interventions will be more effectively delivered. Said delivery will achieve relevant, inclusive and sustainable contributions to humanitarian and development results in an efficient way.* The management domains under which organisational effectiveness will be assessed include strategic, operational, relationship and performance management. Evaluations and results information will be analysed to contribute to the development effectiveness ('results') component of the MOPAN assessments.

MOPAN 3.0 is not an external audit of an organisation, nor is it an institutional evaluation. As such, MOPAN 3.0 cannot comprehensively assess all operations or processes of the organisation. Nor can it provide a definitive picture of the organisation's achievements and performance during the time period (which is often the task of an institutional Annual Report or similar). Nor can MOPAN 3.0 comprehensively document or analyse on-going organisational reform processes.

Instead, MOPAN 3.0 provides a diagnostic of an organisation at a particular point in time, positioning the organisation within its continuum of ongoing improvement. MOPAN 3.0 assessments are discrete; organisations are not compared.

The following series of Operating Principles guides MOPAN 3.0's implementation. This Methodology Manual describes how these principles will be realised.

Box 1: Operating Principles

MOPAN 3.0 will generate credible, fair and accurate assessments through:

- ***Ensuring credibility*** through an impartial, systematic and rigorous approach
- ***Balancing breadth with depth***, adopting an appropriate balance between coverage and depth of information
- ***Prioritising quality*** of information over quantity
- ***Adopting a systematic approach***, including the use of structured tools for enquiry/analysis
- ***Ensuring transparency***, generating an 'audit trail' of findings
- ***Being efficient***, building layers of data, seeking to reduce burdens on organisations
- ***Ensuring utility***, building organisational learning through an iterative process and accessible reporting
- ***Being incisive***, through a focused methodology, which provides concise reporting to *tell the story* of an organisation's current performance

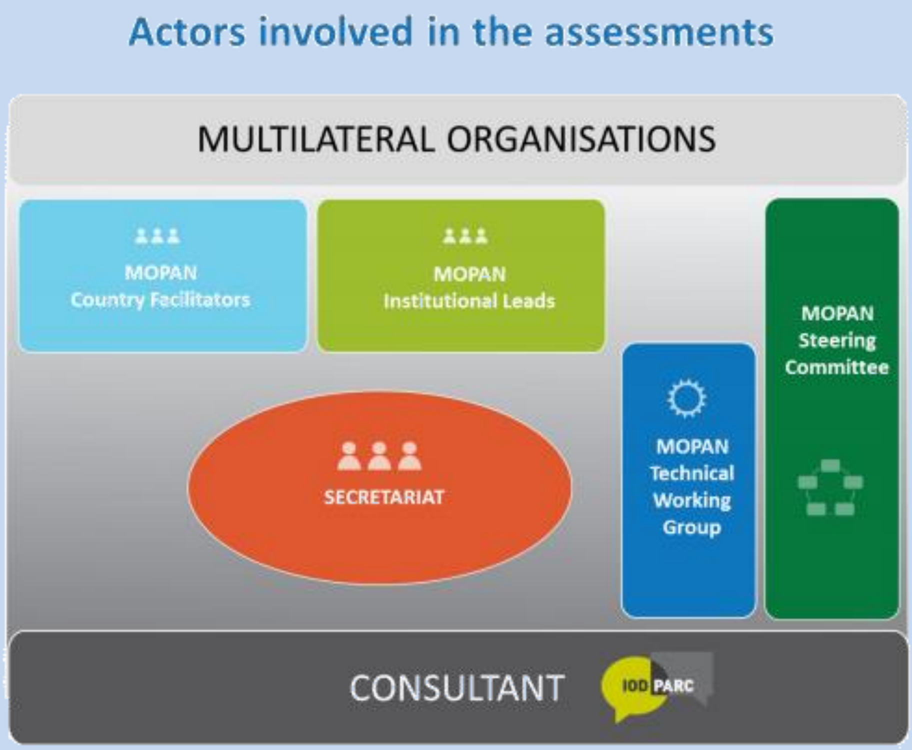
This Manual is intended for use by those interested in the empirical design of MOPAN 3.0 and particularly of the individual component methods.

When reviewing this Methodology Manual, sections of interest may be read, rather than the document requiring reading as a whole. Accordingly, the links are provided from the Table of Contents to the relevant sections.

Actors involved in the assessments

MOPAN 3.0 benefits from close involve of the member governments of the Network, at multiple levels. Actors involved in the Network and the assessments are visualised below.

Figure 1: Actors involved in the assessments



The Steering Committee: The Network overall is governed by a Steering Committee with representatives from all MOPAN member countries. The Committee meets two to three times per year to discuss and decide upon all issues related to the Network, including what organisations to assess and how. The Steering Committee approves all final assessment reports.

The Technical Working Group: The MOPAN Technical Working Group (TWG) is a body of the Steering Committee. Its purpose is to work in partnership with the Secretariat to: provide strategic and technical inputs and oversight in the development and application of the MOPAN approach and methodology for assessing multilateral performance, identify emerging methodological and technical issues, and present recommendations for the approval of the MOPAN Steering Committee.

MOPAN Institutional Leads: MOPAN Institutional Leads represent the network by serving as expert focal points, oversight and champions supporting the assessment process. In their role, they inform MOPAN members familiar with the multilateral organisation under assessment about the process, host the introductory meeting between the multilateral organisation, MOPAN representatives to the organisation and the coordination/research teams. They also provide input to the document review and are consulted on emerging findings prior to finalisation of the institutional report, and present the final report to organisation management and interested board members. Institutional Leads are nominated by their MOPAN Steering Committee representative to serve the network for the duration of one assessment cycle.

The MOPAN Secretariat: The MOPAN Secretariat, hosted by the Organisation for Economic Cooperation and Development (OECD), is the key interface for all parties involved in the process. It is responsible for coordinating and delivering MOPAN assessments.

The service provider conducting the assessments: The independent consulting firm IOD PARC has been appointed by MOPAN to implement the MOPAN 3.0 assessments, with assistance from specialist partners who will support the survey process at country level.

SECTION 2:
SUMMARY DESIGN OF MOPAN 3.0

Achieving the right approach for MOPAN 3.0 depends on ***asking the right questions***. In light of the changing context surrounding multilateral organisations, the following questions have framed the development of MOPAN 3.0.

Box 2: Framing Questions

1. Do multilateral organisations (MOs) have sufficient understanding of the needs and demands they face in the present, and may face in the future?
2. Are MOs using their assets and comparative advantages to maximum effect in the present, and are they prepared for the future?
3. Are their systems, planning and operations fit for purpose? Are they geared in terms of operations to deliver on their mandate?
4. Are MOs delivering and demonstrating relevant and sustainable results in cost-efficient way?

To address these questions, a Theory of Change has been developed. It provides the conceptual framework for the empirical approach for MOPAN 3.0 assessments.

The Theory of Change, and its accompanying narrative, can be found in Section 3 of this Manual. It is based on the following central hypothesis:

Box 3: Hypothesis underlying the MOPAN 3.0 Theory of Change

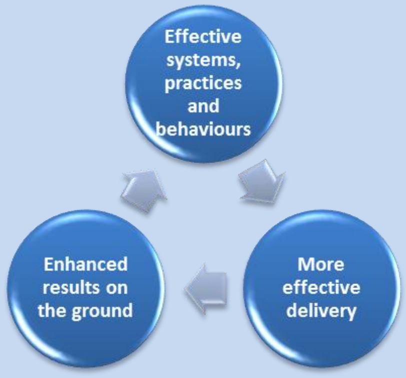
IF a multilateral organisation has effective systems, practices and behaviours in place (*in terms of strategic, operational, relationship and performance management)...*

…..**then** its interventions/activities will be more 'effectively delivered' *(defined as evidence-based, relevant/appropriate, responding to global normative priorities, efficient, functioning within a coherent partnership, with results reported and accounted for)…*

….**and hence,** delivery will achieve relevant, inclusive and sustainable contributions to humanitarian and development results in an efficient way.

This hypothesis can also be articulated as an 'effectiveness loop'.

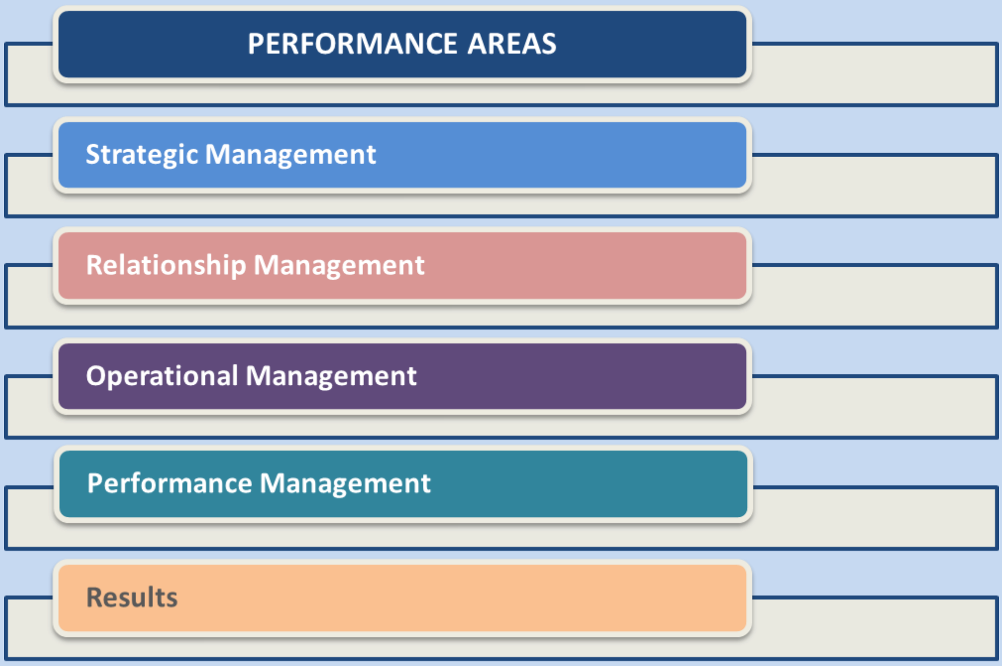
Figure 2: The effectiveness loop



Thus, organisational and development/humanitarian effectiveness are not perceived under MOPAN 3.0 as discrete performance areas, but rather as a **continuum**.

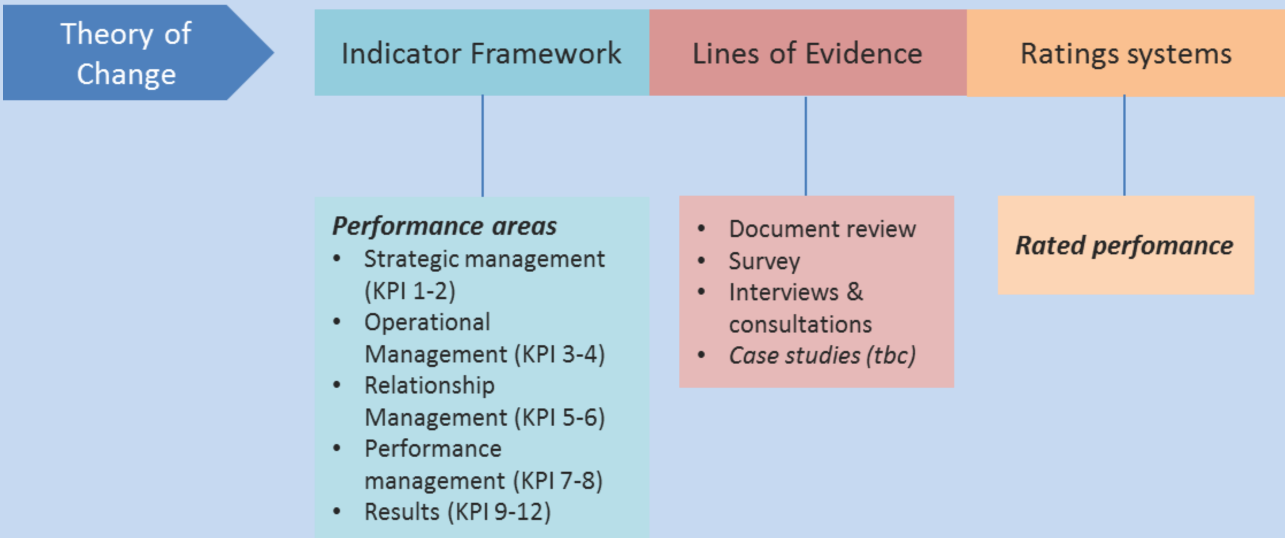
MOPAN 3.0 assesses multilateral organisation performance, across five **Performance Areas**. Four of these areas: **Strategic, Operational, Relationship** and **Performance Management**, relate to aspects of organisational effectiveness. The Performance Area on **Results** addresses development/humanitarian effectiveness.

Figure 3: Performance Areas



The **Empirical Design** of MOPAN 3.0 is grounded in the theoretical approach, and shaped to the Performance Areas. The individual elements of the methodology, and their relation to the Theory of Change, are set out in Figure 4 below.

Figure 4: Empirical Design



The components of this design comprise the following:

i. The **Theory of Change**, which forms the theoretical basis to the assessments. This is presented in Section 3 of this Manual and will be tested and interrogated throughout the assessment process.
ii. An **Indicator Framework**, described in Section 5 and presented in Annex A (with a comparison of current against former indicators is available in Annex B). This presents the twelve Key Performance Indicators and accompanying Micro Indicators against which evidence-based judgments will be made.
iii. **Four lines of evidence**. Four key lines of evidence (methods) are applied: a **document review,** a **survey, interviews** and **consultations**. These are elaborated upon in Section 4 of this Manual. Possible case studies are also under consideration.
iv. The **Scoring and Rating System** provides the system for scoring against individual Micro Indicators and rating against Key Performance Indicators. It is discussed in Section 6 of this Manual.

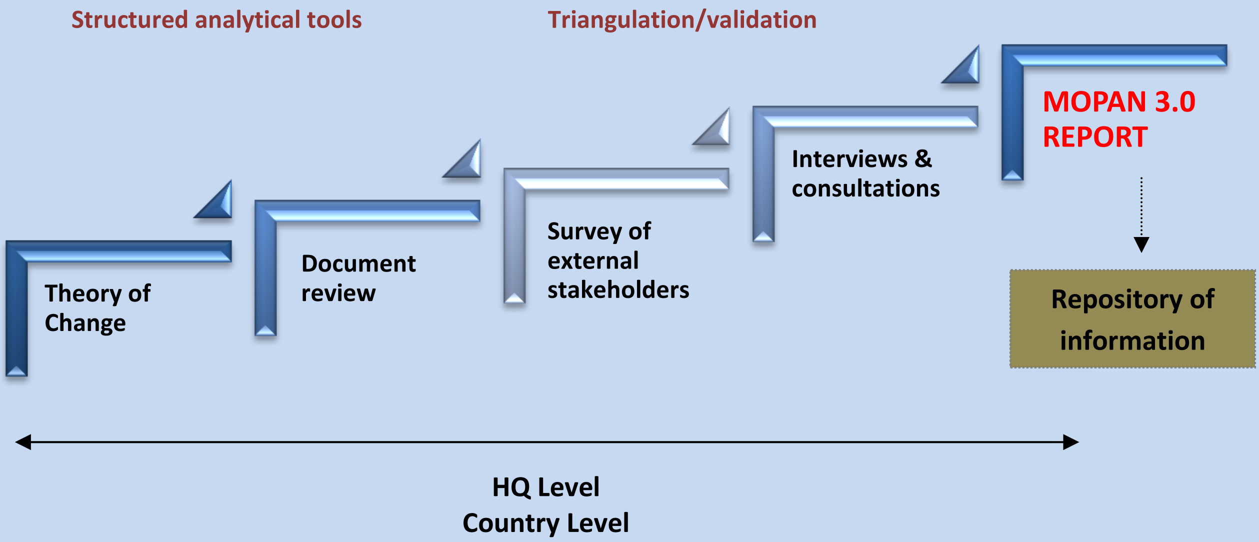
The MOPAN 3.0 methodology applies four specific **approaches:** (1) a sequenced approach, (2) a holistic approach, (3) a systematic approach, and (4) inclusion of country/regional level information.

The approaches will be operationalised as follows:

A **sequenced approach**, will take place through a building blocks model, with each layer of evidence generated through the sequential assessment process, informed by, and building upon, the previous one.

Each **layer of evidence** in the sequence of the assessment process informs the next, as far as feasibility permits. So the survey, for example, will be informed to the best extent possible by findings from the document review. Interviews and consultations will be shaped by findings from both the document review and the survey.

Figure 5: Sequenced approach



A **holistic approach** seeks to apply **multiple lines of evidence** to as many indicators as possible, particularly within Performance Areas on Strategic, Operational, Relationship and Performance Management. The Evidence Density Matrix in Annex C provides a map of intentions here (although these will naturally vary per organisation). This approach will also support triangulation.

Applying **systematic approaches** to both collection and analysis will involve, for data collection, the use of structured frameworks for each evidence line (document review, survey, interviews and consultations). For data analysis, this will entail undertaking structured analysis at aggregate level, applying the different lines of evidence within a composite analytical framework, and applying techniques for validation and triangulation (see Section 7).

Country/regional level evidence will be applied against all relevant indicators, generated through document review and survey particularly. Analysis will incorporate this information throughout, in aggregate form, with relevant examples provided where appropriate. An organisation's country level performance is not scored or rated, and assessments do not contain individual 'country chapters', but rather individual data pieces gathered from country level are reflected in the overall analysis and reporting (see Section 8 on Evidence Management).

The following sections of this Manual describe in detail each of these elements of the methodology.

SECTION 3: THEORY OF CHANGE

This section of the Methodology Manual describes the Theory of Change for MOPAN 3.0. It is based on the central hypothesis presented in Section 2, Box 3, above. To reiterate the main premise:

If a multilateral organisation has effective systems, practices and behaviours in place, **then** its interventions/activities will be more 'effectively delivered', **and hence,** delivery will achieve relevant, inclusive and sustainable contributions to humanitarian and development results in an efficient way.

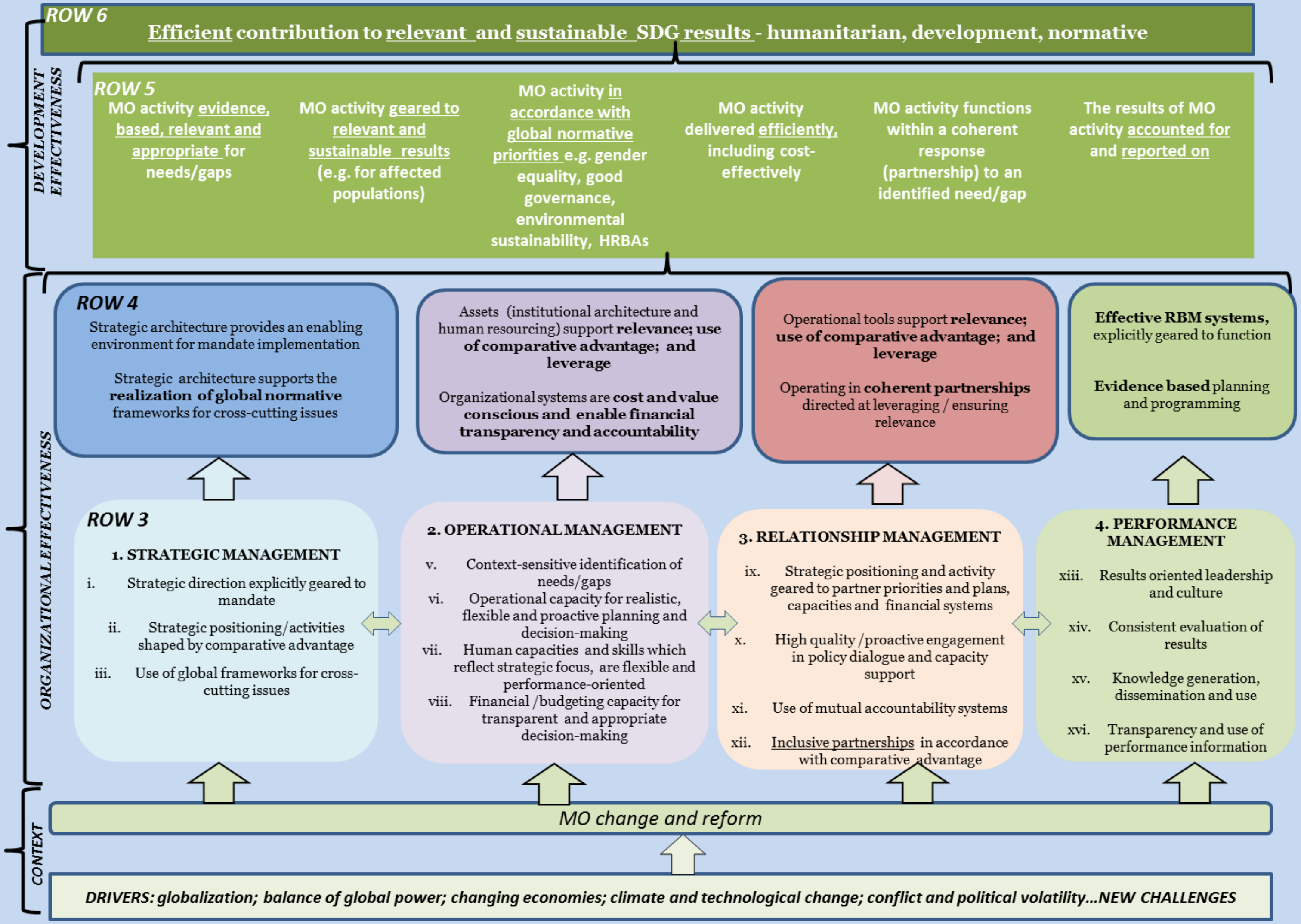
Guided by a range of conceptual, theoretical and empirical literature, the Theory of Change integrates the key international principles, commitments and criteria for humanitarian and development practice. These include:

- The OECD-DAC criteria for development evaluation and those for international humanitarian evaluation[1]
- The principles of the 2005 Paris Declaration on Aid Effectiveness and the 2011 Busan Partnership agreement for Effective Development Co-operation;
- The International Humanitarian Principles.

The Theory of Change is presented below (Figure 6). In summary, its logic (theorised pathways of progression) reflects the understanding of organisational and development effectiveness as a **continuum**.

[1] See ALNAP (2006) Evaluating Humanitarian Action using the OECD DAC criteria

Figure 6: MOPAN 3.0 Theory of Change



The following narrative explicates the Theory of Change for MOPAN 3.0:

Context shapes the policy, strategic and operational needs and priorities confronting multilateral cooperation. Context is reflected in several ways within the theory of change. First, the base layer – reflected in the bottom bar of the schematic – identifies the features of the global environment with implications for multilateral cooperation. Above this, the process of multilateral organisation reform is explicitly identified.

The **logic**, or pathways of progression, for the MOPAN 3.0 theory of change reflects the approach of organisational and development effectiveness as a continuum, with all **pathways connected**. Specific theorised pathways are as follows:

Row 3 posits MOPAN's four Performance Areas of organisational effectiveness (Strategic, Operational, Relationship and Performance)

Rows 4-6 set out, in varying levels of hierarchy, key results (whether humanitarian, development or normative) which can, according to the theory, be reasonable expected at different levels. Thus:

Row 4 provides the key results in terms of organisational effectiveness which are posited by MOPAN (and reflected in the Key Performance Indicators, discussed in Section 5 and provided at Annex 1). These offer MOPAN 3.0's response to the question: *What does an effective multilateral organisation look like in terms of its systems, practices and behaviours?* It will be assessed through analysis against the Performance Areas on Strategic, Operational, Relationship and Performance Management.

Row 5 marks the transition into development effectiveness. It reflects the explanation above of effective delivery, or "the characteristics of multilateral organisation programmes and activities which are clearly linked to supporting the achievement of results."[2] It will be assessed through a combination of all five Performance Areas.

Row 6 is the highest level of the hierarchy, and indicates the range of results which may be achieved, formulated as contributions – whether normative, humanitarian or development – to global frameworks (e.g. Agenda 2030). It will be assessed through analysis against the Performance Areas on Results.

Assumptions

Within any Theory of Change, assumptions explain the connections between early, intermediate and long-term changes, and the expectations about how and why proposed interventions will bring them about. Often, assumptions are supported by empirical evidence, strengthening the case to be made about the plausibility of theory and the likelihood that stated goals – here, contributions to Agenda 2030 – will be accomplished.

[2] OECD DAC Development Effectiveness Review

The process of developing the Theory of Change has brought to light several key assumptions, some of which are also highlighted in the 2013 Independent Evaluation of MOPAN.[3] These have been tested through the assessment process, and revisited following the completion of the first round of assessments.

They have been separated into principal and secondary assumptions, with two principal assumptions relating to the overall theory of change, and the latter linked to the progression upwards through the theory of change. Assumptions per Performance Area have not been treated, since these would be excessively numerous and detailed. More assumptions will become apparent as the assessment process for Cycle 2 proceeds: those presented here are indicative rather than exhaustive.

Principal Assumptions

- That organisational effectiveness has an influence on an organisation's ability to deliver its programming in more effective ways, and thus achieve its strategic objectives and contribute to its proposed development or humanitarian results (Rows 4-6)
- That improving, in differing combinations, an organisation's strategic, operational, relationship and performance management will contribute to its organisational effectiveness (Rows 4-5)

Secondary Assumptions

From context to organisational effectiveness

- That the (live) organisational reform/change strategy and action plan of the multilateral organisation is aligned to and *in sync* with the relevant Performance Areas and the overall [MOPAN] vision of an effective multilateral organisation

From facets of organisational effectiveness [row 3] to [row 4] to expected characteristics of activity

- That the organisation has a clearly articulated and consistently held view on its comparative advantage *[strategic management]*
- That the organisation has a sufficiently stable governance and financial environment in which to run its operational management systems *[operational management]*
- That the organisation has a clear understanding of its rationale/its approach to partnerships and a distinction between the different types of partnerships in which it participates *[relationship management]*
- That the organisation has a clear and consistent position, reflected through the different strategy/programming levels of the organisation, on the nature of its targeted and tracked results (outputs, outcomes, impact) *[performance management]*

From organisational effectiveness [row 4] to [row 5] development effectiveness

- That the organisation operates within a cycle of strategy setting to programming/work planning that allows for visibility of its intended organisational effectiveness within its results
- That the organisation is balancing its responsiveness to new agendas and opportunities with its existing programming against its core mandate

[3] Independent External Evaluation of MOPAN. Balogun, P et al. (2013)

- That the organisation has an established and consistent view on the parameters by which it judges cost-effectiveness

These assumptions will be tested and reported on through the assessment process; and others will be identified – and also tested – within individual assessments.

SECTION 4: INDICATOR FRAMEWORK

MOPAN 3.0 applies an Indicator Framework aligned to the five Performance Areas set out in Section 2 above. The Indicator Framework is comprised of eight Key Performance Indicators (KPIs 1-8) aligned to the four Performance Areas on Organisational Effectiveness (strategic, operational, relationship and performance), and four Key Performance Indicators (KPIs 9-12) aligned to the Performance Area on Results.

The generic indicator framework for MOPAN 3.0, available in Annex 1, is purposefully:

- **Geared to the Theory of Change**, reflecting updated concepts and 'continuum' or organisational and development effectiveness
- **Aligned to all five Performance Areas**
- **A generic model**, subject to adaptation, below, the indicators are designed to be applicable to varying types of multilateral organisation, and varying types of activity/intervention
- Built to enact the principle of **'function' over 'form'**, with the presence of a system, behaviour or practice not sufficient to equate to evidence of 'effectiveness'(which is rather reflected in its application in practice

Adaptation/tailoring of indicators

Whilst the five Performance Areas are broadly applicable to the range of multilateral organisations to be assessed under MOPAN 3.0, the dimensions explored have previously been adjusted, to reflect the mandates and maturity of each organisation.

Customisation has, in the past, been handled on a case-by-case basis between organisations under assessment and assessment teams. As per the Operating Principles of producing ***credible, fair and*** **accurate** assessments, the approach for adaptation will be operationalised as follows:

1) Indicators in the revised framework have been ***formulated to explicitly adopt a generic model. Therefore, all or most indicators should apply to organisations under assessment, unless there is a clear evidence-based case of non-applicability.*** This applies to whether assessed organisations are UN agencies, International Financial Institutions or Global Funds, and whether they are undertaking normative, humanitarian or development work.

2) As adaptation and tailoring under MOPAN 3.0 will be lesser in extent, indicators will be applied ***as relevant to the organisation's mandate and operating practice.*** For example, development results will be assessed according to the organisation's own corporate results framework.[4]

Process wise, this will encompass the steps outlined in Table 1, below.

Table 1: Steps for adaptation/tailoring of indicators

Step 1	The operating principle above will apply, of the application of indicators *in their current formulation* unless there is a clear, evidence-based case of non-applicability

[4] With commentary to be supplied on the quality of the corporate results framework and the data sources supplying it – see Generic Indicator Framework, Annex 1

Step 2	Where there is a clear case that the indicator does not apply to the organisation, i.e. is not part of their mandate or operating model, the indicator will be omitted For example: *MI 3.3: Aid reallocation/programming decisions can be made at a decentralised level under delegated authority within an appropriate budget cap* will not apply to an agency with no decentralised structure and which does not conduct programming (e.g. the Global Fund for Aid TB and Malaria, which has specific operating structures)
Step 3	Where the indicator can be applied, but requires adaptation for organisational relevance, it will remain unaltered, but a differential interpretation will be agreed and made explicit.

Two key nuances to this approach are as follows:

1) For cross-cutting issues: these are universal development and humanitarian aims, legitimised and mandated by global frameworks such as Agenda 2030. Many multilateral organisations work on these issues, although they may not feature explicitly in mandates. These will therefore be assessed **where there is a clear statement of intent by the organisation to take these issues into account in their work** (e.g. at strategy level). Other than where they are, for example, treated as system wide goals,[5] assessments will not seek to apply an externalised ‘benchmark’ or ‘standard’ on which to assess performance. Rather, they will be interpreted **as applied within a multilateral organisation’s operating model, business practice and results**.

For agencies whose mandates explicitly target a specific cross-cutting issue, e.g. gender, this will not for a focus of assessment, since it is a given in terms of thematic interest. However, their treatment of the issue, through the systems, practices and behaviours applied for organisational effectiveness, and the results they achieve – will be assessed through the normal process.

2) A maximum of five additional organisation-specific indicators can be considered. However, this must be subject to agreement from the MOPAN Institutional Lead and from the concerned organisation.

[5] Reflected in some cases in system-wide instruments and internal assessments such as the UN System Wide Action Plan for Gender Equality and Women’s Empowerment

SECTION 5: EVIDENCE STREAMS

Beneath the theoretical framework of MOPAN 3.0 sits the empirical design set out in Section 2 above. To operationalise the empirical design, and as set out in Figure 4, four **evidence streams** will be applied: a document review, a survey, interviews and consultations. These evidence streams are described in this Section of the Manual.

5.1 Document Review

The MOPAN Document Review provides a key vehicle to explore 'evidence that multilateral organisations have the systems, practices, or behaviours in place that MOPAN considers to be important factors in an organisation's effectiveness and evidence of its contributions to development and/or humanitarian results.'[6] The Document Reviews are a key plank in the MOPAN methodology, forming the basis on which other evidence streams will rest.

Purpose of the Document Review

Through an examination of relevant documents, the Document Review component of MOPAN 3.0 explores two areas:

1) Whether multilateral organisations have the **systems, practices, and/or behaviours in place** that MOPAN considers to be important dimensions of organisational effectiveness, as reflected in the indicator framework

2) Particularly through a combination of evaluations, management results and other evidence, **evidence of contributions to development, normative or humanitarian results** (development/humanitarian effectiveness)

Additionally, and under the sequenced approach above, evidence from the Document Review provides the starting evidence block to inform more focused enquiry under the survey and the interviews/consultations. It also plays a major role in triangulation.

Approach

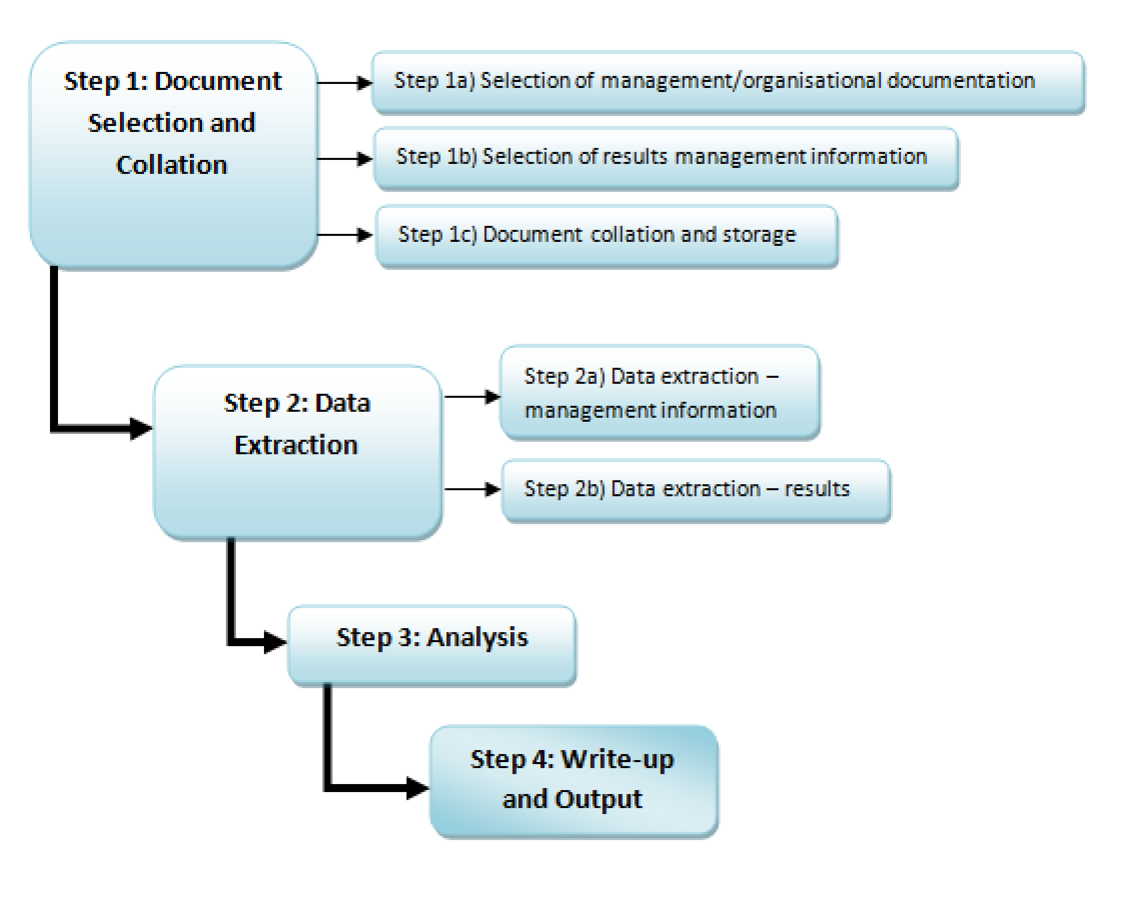
The Document Review seeks evidence against the **individual Micro Indicators** of the indicator framework (see Section 4 and Annex A). This in turn supports the development of the narrative against the KPIs.

Step-by-step Approach

The Document Review, combining management information and independent evidence, will adopt the following process. The individual steps and sub-steps are set out in the subsequent text.

[6] Terms of Reference

Figure 7: Structure of the document review



Step 1: Document Selection and Collation

a) Selection of Management/Organisational Documentation

Given the wide range of organisations that MOPAN assesses, there is no *once size fits all* approach to designing an appropriate selection of documents. However, under the Operating Principle of ***efficiency***, burdens on assessed organisations for document supply need to be reduced. Under the Operating Principle of ***fairness***, similar types of documentation need to be analysed per organisation. The volume of assessments under MOPAN 3.0, require an emphasis on ***feasibility.***

Document reviews will apply **purposive selection**, aiming to provide coverage of all relevant Micro Indicators, adapted if appropriate to the organisation, guided by a clear and consistent typology (see Section 5). To meet both feasibility and coverage concerns, a selection of documents per organisation will be collated, drawn from a wider screening process, to provide a depth of insight into the organisation's recent and current status, up to a ceiling of approximately 60 documents. Additionally, up to 20 evaluations will be selected (if available).

The following typology will be applied:

Box 4: Document Review Indicative Typology

i) **External assessments**[7], such as:

- UN Joint Inspection Unit reports
- Previous MOPAN Assessments
- Quality of Official Development Assistance conducted by the Brookings Institution and the Center for Global Development
- OECD DAC Development Effectiveness Reviews
- Major institutional evaluations
- Peer Reviews of Evaluation Functions
- Other independent or externally conducted reviews or studies of performance

ii) **Evaluations**, including:

- Independent evaluations
- Internal evaluations
- Joint evaluations

iii) **Management information**, such as:

- ***Corporate strategic plans***, results framework and reporting processes (Annual Report or similar)
- ***Regional strategic plans,*** results frameworks and reporting processes (Annual Regional Report or similar)
- ***Selection of policies,*** their results frameworks and reporting processes
- ***Selection of sector strategies/plans***, results frameworks and reporting processes
- ***Selection of country strategic plans (sample countries),*** results frameworks and country reporting
- ***Institutional architecture information***
- ***Key corporate financial and budget information*** (Five-year Financial Framework/biennial budgets and reports, Annual Financial Report, etc.)
- ***External audits***
- ***Key business process documentation*** which relates to relevant MIs (Performance Management systems, Human Resource Management Strategy, Resource Mobilization Strategy, financial control mechanisms, e.g. internal audit strategy, risk management strategy, anti-corruption strategy, programme design and approval documentation, social safeguard systems, evaluation quality assurance processes, etc.)
- ***Executive Board minutes and decisions***

[7] The review of documents excludes bilateral assessments of the multilateral organisations.

The parameters for individual document status are defined in Table 2 below:

Table 2: Parameters for Document Status

Legitimacy	• All documents must be in final form to be included in the assessment • All documents must be recognised by the institution's management
Accessibility	• Publicly available documentation will be used (information sourced from web pages) • Where this is not directly available, organisations will be approached directly following the protocols set out in Document Collation and Storage, below
Timing	• Policies or guidelines, at any level within the multilateral organisation, are selected only if they are in force as of the year the MOPAN assessment of that organisations begins • Strategies, regardless of level within the multilateral organisation, are selected only if they are being implemented during the period the MOPAN assessment covers • Any information presented on the multilateral organisation's website will be retrieved within the period the assessment covers and is assumed to be current unless the web page itself states otherwise • All documents (except for policies, guidelines and strategies) should be published within three years prior to the start of the assessment process[8], unless there is a strong rationale for reviewing older documents: ➢ Project/programme level documents ➢ Country, regional or organisation-wide documents: from three years prior to the start of the assessment ➢ Evaluations: from two years prior to the start of the assessment
Sub-category selection	• Analysis of relevant regional strategies and results frameworks • Country level documentation, three types of information: ➢ Country strategy and results frameworks ➢ Country strategy evaluations, if available, from up to two years prior to the start of the assessment ➢ If country strategy evaluation is unavailable, then one other key strategic document at country level • When specific MIs require a selection of sector strategies, country strategies, or project level documentation, a specific selection approach should be developed and tailored for each multilateral organisation

b) To assess progress towards results, the results management system of the organisation will be assessed as part of the Document Review. This will include:

- review of management results produced, and whether these apply recognised international data sources/are triangulated/ are independently verified
- review of the evaluation function to consider the independence and impartiality of evaluations produced

[8] A three year indicative timeframe for relevant documentation will be applied, though with flexibility where appropriate, noting that strategic planning and evaluation cycles may vary.

This approach will produce different types and levels of results information, as follows:

Box 5: Sources of results information

Management results	Evaluations or reviews
i. Management results which rely on internally-generated data ii. Management results which apply recognised national or international data sources (e.g. international health statistics) iii. Management results which apply recognised international data sources (e.g. international health statistics) and are independently verified	i. Evaluations or reviews conducted internally ii. Evaluations which are conducted independently iii. Evaluations which are conducted independently and which undergo formal quality assurance processes

Results evidence from both evaluations and management information will be included in analysis against KPIs 9-12. The potential range of sources is below:

Box 6: Documents regarding results information

- **Annual Report on Development Results** at the organisational level. For International Financial Institutions (IFIs) this is often based on a review of project completion reports which may or may not have been audited for accuracy by the central evaluation group. For United Nations (UN) organisations this report may track SDG results across partner countries and may be supplemented by highlights from evaluations. For UN agencies it can usually be found in the documents submitted to the governing body at its main annual meeting.
- **An Annual Summary/Synthesis of Evaluation Results**. This is a common document among IFIs and some UN organisations and typically presents both extracted highlights and some statistical data on the coverage and results of evaluations published in a given year.
- **OECD Development Effectiveness Review**, which for some multilateral agencies provides a synthesis of development results, drawn from evaluations, for the organisation.
- **A Report on Progress towards the Objectives of the Strategic Plan**. This report is not necessarily issued annually. It can either relate to the biennial budget of an organisation or to a three to five year strategic plan of an IFI. It may be in a Scorecard or Results Framework report format.
- **Independent evaluations** commissioned by the organisation including a selection of:
    - Thematic evaluations
    - Strategic evaluations
    - Country/Regional evaluations
    - Programme evaluation
    - Any other relevant evaluation type
- **Evaluations or reviews conducted of the organisation and which include a summation of its results** (for example evaluations or reviews commissioned by bilateral donors)

The validity and credibility of the results evidence is a key issue in relation to evidence of results against KPIs 9-12. For example, evidence from independently-conducted evaluations, which have undergone a rigorous quality assurance process, holds a different level of credibility from internally-generated management results which are not triangulated or independently verified. Early review, as part of the Interim Document Review, on how an organisation's results evidence is constituted will allow for greater clarity at analysis stage, and will be transparently presented in assessment reports (with clear sourcing provided).

The selection of evaluations (where relevant) to provide evidence on results does not **aim to be representative in terms of coverage** (e.g. of an organisation's financial expenditure per year). Rather a purposive selection will take place, based on agreed parameters.

The selection process for selecting individual evaluations will involve the sub-steps laid out in Table 3 below.[9]

Table 3: Sub-steps for selection of evaluations

Sub-step	Action
1	Examine a subset of the available evaluation reports to determine how their scope of coverage is defined, e.g. in terms of geographic coverage, thematic coverage, objectives coverage, policy areas, technical focus etc.
2	Agree the primary measure(s) to assess coverage of the given selection of evaluation reports. This will vary according to how the organisation arranges its evaluation coverage but in all cases will aim at maximising breadth of coverage within the resourcing available. Examples include: • Strategic importance (e.g. major evaluations of key thematic areas of engagement by the organisation) • Geographic coverage (e.g. a selection of evaluation reports in the three regions of highest investment by the multilateral organisation) • Policy/sector coverage (e.g. a selection of evaluation reports of the four main policy areas/sectors of investment by the multilateral organisation) • Strategic objective coverage (e.g. a selection of evaluations that address the key strategic objectives of the organisation) • Other key priority thematic or technical areas such as gender, climate change, etc.
3	Purposively select up to 20 individual evaluation reports using the criteria defined above, aiming for maximum strategic importance and diversity.

c) Document collation and storage

For all elements of the Document Review, **the protocols below will apply to document collation and storage:**

- Identification of the specific documents to be analysed according to the typology above. This will take place through an initial web search for relevant documents available publicly on the websites of the multilateral organisations, and, for example, OECD-DAC.
- From a wider screening process, documents will be extracted and sourced in a 'library'. A list of the specific documents for analysis will be prepared;
- Consultation will then take place with the organisation, via the Secretariat (with inputs from the Institutional Lead(s)), to ensure that the list is complete, and that any additional documents (such as Executive Board Minutes) can be supplied by the organisation.
- Additional documents will be sourced at the stage of updating the Document Review in a second stage

[9] Adapted from the DER sampling process

- Documents gathered for the assessment will be stored in a storage system that can be accessed across the internal assessment team and will be filed systematically according to organisation and relevant document type.

A full Bibliography of all relevant sources (as constitute the final document set used) will be prepared for discussion/agreement with the organisation undergoing assessment.

Step 2: Data Extraction

Data extraction entails identifying and extracting evidence aligned to the relevant MIs, without (as yet) applying a layer of analysis. This is the first step in ensuring a clear evidence trail from data to findings.

Data extraction will take place through the application of a ***structured analytical tool***, geared to the MIs, to ensure systematic capture and analysis of data across the main analytical fields of the study (here, the KPIs, MIs and their associated criteria). Extraction will take two forms, depending on the information source.

a) Data extraction of management information

Systematic extraction of data from the selected documents will take place against the analytical template. This will involve plotting in data (and, critically, sources) against the identified analytical fields (MIs) and criteria. Data gaps will need to be explicitly flagged for later mitigation, if possible, through other evidence lines.

Extraction will take place against the relevant MI. **It is important that no judgment is formed at this stage.** To enhance substantive findings, additional information against five parameters will also be sought:

ii) Consistency/improvement over time, whether e.g. a policy or initiative has been developed and implemented, and whether and how it has led to changes in practice/improvement in results

iii) 'Evidence of implementation,' to ensure a movement beyond 'form' to whether policies and strategies are being implemented in practice

iv) Context, key features of the organisation's operating environment identified to both provide explanatory factors and to populate the specific theory of change for the organisational assessment

v) Innovation, any evidence of innovation (of different types)[10]

Data against the relevant MI will be plotted into the analytical template as it appears in the source document and clear referencing (document number and page reference) provided. A bibliography will be prepared which includes the full title, date of publication, etc. (applying the OECD's referencing system).

b) Data extraction on results

Once the form of results information has been clarified, and the sample of evaluations produced, data will be extracted from the evidence against the analytical template. The *status* of the results information, applying the hierarchy above, will be clarified.

[10] Evidence of innovation will be broadly captured at this stage, since there is no agreed definition of innovation, and understanding of it can vary greatly in perception. Therefore, it is intended to capture evidence at this stage, which may be subject to analysis later, e.g. categorisation of 'types' of innovation

Step 3: Analysis

Once evidence has been extracted and plotted into the analytical template, analysis will then need to take place.

As one evidence stream contributing to overall analysis, Document Reviews will not be scored. Nonetheless, it is important that a layer of analysis is applied to the Document Review to provide a substantive foundation upon which the successive evidence streams (surveys, interviews and consultations) can be based.

Analysis for the Document Review will have the following characteristics:

- Triangulation will take place by mapping data sources per MI against each other in the analytical matrix, and flagging any apparent tensions or contradictions – to provide any opportunities for later deepening or interrogation through survey or interview.
- Analysis will identify key themes and common densities of evidence against the individual MIs. It will then describe findings according to the evidence available.
- Analysis will be thematic in nature, describing the frequency, intensity and significance of findings, and be explicit on the strength of evidence in particular areas.
- Emphasis will also be placed on (and space provided in the framework for) explaining the reasons/explanations for good or under-performance
- For each MI, a narrative will be generated which summarises the evidence against it, presenting the key themes arising from the evidence. It will report this in broad terms of **'strong evidence of [good performance/underperformance]'** or **'weak/insufficient evidence – requires further assessment through other evidence streams.'**

Step 4: Drafting and updating of the Document Review

a) An Interim Document Review will be produced for each multilateral organisation during the first year of the assessment process. The Interim Document Review reports on each MI with a summary under each KPI, linked to the relevant documentary sources, with clear traceability to sources. The multilateral organisation and the Institutional Lead will have the opportunity to comment on the Document Review.

b) Following comments on the Interim Document Review, a subsequent round of data extraction and analysis will take place at an agreed point later during the assessment process. This will ensure that the assessment captures all relevant documentation up to an agreed cut-off point, as well as integrating comments received on the Interim version. The final Document Review will constitute one of the four evidence streams to input into the final analytical process for the assessment.

Standardisation/Consistency

Standardisation and consistency will be achieved through the following steps:

- Training will take place of researchers dedicated to the Document Review process, in order that all have a full understanding of the indicator framework and the Scoring and Rating system

- For each Document Review, the research analyst assigned to each organisation will conduct the initial data extraction. One key document and the relevant data extraction will then be peer reviewed by the Core Team Member designated to oversee the assessment process for that specific organisation;
- For analysis, a similar process will be adopted, where the analysis against a sample set of MI will be reviewed by the designated Core Team Member, and amended as appropriate. Consistency rating will also take place among the three Core Team Members;
- To ensure consistency, assessment teams for each organisation will hold regular joint meetings to ensure common interpretation of the various indicators, to note data sources, and standardise as far as possible the data analysis process;
- Finally, the resulting draft Document Review will be quality assured by the designated Core Team member.

5.2 Survey

The survey seeks to gather both perception data and an understanding of practice from a diverse set of stakeholders of the multilateral organisation under assessment; for example, whether respondents consider the organisation has a clear strategic vision in place; or whether a specific policy, strategy or business practice has permeated to country-level. As well as a substantive dataset in itself, therefore, the survey performs a key element of triangulation and verification for the other data collection methods.

The survey will be applied online in the first instance, though off-line methods may be used as a contingency. Respondents may fill out a paper-based survey, complete an electronic version of the survey in Microsoft Word that is sent be email or, in exceptional cases, participate in a structured interview either in person or by telephone. Respondents will be able to complete the web-based survey in English, French or Spanish, as well as specific local languages which change each cycle given the selection of countries on an ad-hoc basis. A professional survey company will provide support to the external service providers in the design, implementation and analysis of the results

Box 7: Learning from previous experience of the survey

In 2013, MOPAN commissioned an independent, external evaluation of the previous methodology, the Common Approach.[1] On the basis of that evaluation and MOPAN member needs, certain changes were introduced in 2014. Nonetheless, a number of concerns appear to remain, these include:

1) Quality of data

- A significant percentage of 'don't know' responses suggest that the level of familiarity of some respondents with the organisation being assessed may not be sufficient to respond to all questions (e.g. on the internal operations of the organisation). This undermines validity.
- The occurrence of respondent fatigue as the survey instrument remained quite long. This may affect both the quality and rate of response.

2) Rating scales

- The rating choices provided in the survey may not be used consistently by all respondents, especially across different cultures involved in the MOPAN assessment. One potential limitation is 'central tendency bias' (i.e. a tendency in respondents to avoid extremes on a scale). Cultural differences may also contribute to this bias.

Respondent types

The approach to the survey under the design of MOPAN 3.0 is to *ask the right questions to the right people.* As such, the MOPAN 3.0 survey will target the following primary respondent groups:[11]

Table 4: Survey respondent types

Location of Respondent	Description of Respondent (desired)
Donor Headquarters Oversight (HQ)	Professional staff, working for a MOPAN donor government, who share responsibility for overseeing/observing a multilateral organisation at the institutional level. These respondents may be based at the permanent mission of the multilateral organisation or in the donor capital.
Donor Country Office Oversight (CO)	Individuals who work for a MOPAN donor government and are in a position that shares responsibility for overseeing/observing a multilateral organisation at the country level.

[11] The number and type of respondent groups may vary for each organisation and additional respondent types may be included.

Direct Partner/Client (DP)	For most of the assessments, this group will include individuals who work for a national partner organisation (government or civil society) in a developing country; or in some cases at the headquarters of the multilateral organisations. Respondents are usually professional staff from organisations that receive some sort of direct transfer from the multilateral organisation or that have direct interaction with the organisation at country level (this could take the form of financial assistance, technical assistance, policy advice, equipment, supplies, etc.). The exact definition of "direct partner" varies according to the context of each organisation assessed. In some cases, direct partners include staff members from international agencies that are implementing projects in conjunction with the multilateral organisation being reviewed.

For some organisations, other respondent categories will also be used, such as peer organisations, co-sponsoring agencies, technical partners and/or recipient/host governments. These respondents will be identified and sourced with support from the relevant contact from MOPAN and/or the MOs; though, the survey team will retain responsibility for maintaining the respondent list. The total number of respondents will likely vary between assessments.

A common threshold of respondents will be set, at approximately 150 per organisation.

Respondent selection – first filter

MOPAN members and the institutions being assessed will be key contributors to the selection of potential respondents, with an emphasis on people who are familiar and knowledgeable about the organisation being assessed. This presents the first *filter* of respondent selection; though it is acknowledged and noted that this is a partial/subjective filter as there are limited means of determining whether the most familiar and knowledgeable individuals are selected.

Criteria for this *first filter* may include:

- **Longevity of engagement**, for example: How long has the respondent been working/ interacting with [the organisation being assessed]?
  - Less than a year
  - At least a year but less than two
  - Two years or more

- **Frequency of engagement,** for example: How often does the respondent have contact with [the organisation being assessed]?
  - Daily
  - Weekly
  - Monthly
  - A few times per year or less
  - Never

Respondent Self-Assessment – second filter

Given the principle of '*asking the right questions to the right people*' it is proposed to raise the threshold level of 'expertise' required to complete the survey denoted by the scale of familiarity (1 – 5; where 1 denotes a 'not at all familiar' level to 5 which denotes a 'very familiar' level). The shift proposes to apply a familiarity rating of Moderately Familiar or above; rather than the former Common Approach method of using the second point on the rating and above threshold. Criteria for this second filter may include:

- **Level of familiarity,** for example: How familiar are you with [the organisation being assessed]?
    - Not at all familiar
    - Slightly familiar
    - Moderately Familiar
    - Very familiar
    - Extremely familiar

The reason for this choice is that the nature of the survey questions necessitates a reasonably high level of understanding and familiarity of the organisation - usually, having direct partnership or implementation experience. Demonstrating the inclusion of an 'informed and familiar' respondent population is also important for the credibility and validity of the survey. This change also responds to feedback presented in the management responses of some organisations to the assessment.[12]

Survey design and customisation

A core set of questions will be developed for all respondents, to ensure consistency, in addition to a level of customisation for specific respondent group questions. A more streamlined set of questions than previously (35-40 questions in total) will seek respondent ratings and qualitative responses against the relevant Micro Indicators of the MOPAN 3.0 indicator framework.

The scope for adaptation/customisation of the survey questions will be explored to better reflect organisations' mandates, operational challenges and reform agendas. This will be done in consultation with the organisations being assessed and other individuals (MOPAN members and external resources) familiar with them.

Some questions may also be specified for specific respondent groups (reflecting their functional responsibility or relationship with the organisations).

Blending Quantitative and Qualitative questions

The central element of the survey consists of a series of closed-ended questions geared to the indicator framework for MOPAN 3.0. The survey will be designed using a rating scale where respondents are requested to specify their rating of performance on a symmetric strong-weak scale to a series of statements. This rating scale is a bipolar scaling method, measuring either positive or negative responses to a statement. An even-point scale will be used to avoid respondents giving a non-committal middle rating. This is sometimes called a "forced choice" method, since the neutral option is removed. A 'no opinion'/ 'don't know' option will be given. In addition to requesting a selection on the rating scale, provision will also be made available to capture qualitative feedback through open-ended questions.

The survey will conclude with respondents being asked further open-ended questions on their views of the MO's overall strengths and areas for improvement.

---

[12] e.g. AfDB MOPAN Assessment 2014 Management Response

Administering the survey

The Service provider will ensure the efficient deployment and overall management of the survey instrument. The initial invitation to participate will be sent from the survey provider, to names provided by MOPAN members and multilateral organisations at country level, with an embedded web-link within the email.

Ethical Standards

The survey will be conducted in adherence to international ethical norms and standards, such as those of UNEG[13] and the OECD DAC. Specifically: to ensure confidentiality and anonymity of respondents, the management and deployment of the survey and conduct of other contingency methods of data collection for the survey will be hosted on secured servers with restricted accessibility; confidentiality will be protected through restricted access to raw data and survey respondent data will be anonymised and themed at the aggregate level.

Transparency

To support independence, but whilst mindful of confidentiality concerns, above, the survey team will adopt a fully transparent approach, presenting content and analytical methods transparently and documenting responses to comments received.

Response rates

Expert opinion differs on acceptable survey response rates (defined as 'the result of dividing the number of people who answered the survey by the total number of people in the sample whom were eligible to participate'14). Based on a recent meta-analysis, 15 of 45 studies examining differences in the response rate between web surveys and other survey modes, it is estimated that the response rate in the web survey on average is approximately 11% lower than that of other survey modes. A low response rate can give rise to sampling bias if the non-response is unequal among the respondents. Such bias is known as non-response bias (see Survey Design, below, for how this will be mitigated).

The revised Common Approach methodology aims at a 70% response rate from donors at headquarters and a 50% response rate from all other target groups. The survey team supports this aim, and it has been adopted for MOPAN 3.0, mindful that a high response rate from a smaller sample is considered preferable to a low response rate from a large sample[16].

During the survey period, response rates will be monitored regularly through the on-line analysis platform by the survey team. Reminders will be issued to those respondents who do not access the survey or who do not complete it from a range of sources:

- MOPAN member country office and headquarter respondents will receive reminders from their MOPAN Focal Point.

[13] UNEG - Ethical Guidelines for Evaluation (2008); UNEG Code of Conduct for Evaluation in the UN system
[14] "Response Rates – An Overview." American Association for Public Opinion Research (AAPOR). 29 Sept 2008. http://www.aapor.org/responseratesanoverview
[15] Web Surveys versus Other Survey Modes – A Meta-Analysis Comparing Response Rates Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., Vehovar, V. International Journal of Market Research, 50, 1, pp. 79-104, 2008
[16] Evans SJ. Good surveys guide. BMJ. 1991 Feb 9; 302(6772):302-3. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1669002/pdf/bmj00112-0008.pdf

- Direct partners and any other respondent groups will receive reminders online, issued and tracked by the survey team.

Regular, but light touch contact will be maintained with the relevant MOPAN contact points throughout the survey period to ensure up-to-date information and timely encouragement to complete the survey is available.

Encouraging Respondents; increasing response rates

The following principles and practices will be mainstreamed into the survey to help increase response rates:

- **Early notification**: an email notification will be sent, notifying participants that they will be receiving the survey
- **Clarity of survey purpose**: participants will be informed of the purpose of the survey and how their feedback will be used
- **Active consideration of respondents' time**: the survey will present an indicative amount of time the survey will take to complete (based on piloting and testing data)
- **Senior 'sponsorship'**: ensuring the email invitation is issued from someone with sufficient seniority that the invite carries importance
- **Accessibility and succinctness**: the survey will be shortened from previous versions to ensure usability, ease of comprehension (no jargon that respondents can't understand) and improve response rates; mitigate respondent burden and survey fatigue; the survey will not be overburdened with unnecessary questions
- **Showing progress**: a progress bar (or other notification) will indicate how much longer the survey will take
- **Proportionate number of reminders**: two reminders will be issued at milestone points in the survey completion window to those that have not completed the survey
- **Optimisation of survey for all devices**: where possible, surveys will be optimised for completion on devices from desktop PCs to mobile devices

Mitigating bias

Bias can be placed under two categories; nonresponse bias and response bias.

Nonresponse bias occurs when some respondents included in the sample do not respond; that is, the error comes from an absence of respondents instead of the collection of erroneous data. Therefore, nonresponse bias is the variation between the true mean values of the original sample list (people who are sent survey invites) and the true mean values of the net sample (actual respondents). Most often, this form of bias is created by refusals to participate or the inability to reach some respondents. To mitigate nonresponse bias, the following actions will be taken:

- **Thorough pre-test of survey mediums** – the survey and invitations will be piloted and pre-tested to ensure they run smoothly through any medium or on any device potential respondents might use. Through an acknowledgement of the respondent sample's different forms of communication software and devices and pre-testing the surveys and invites on each, this will ensure the survey runs smoothly for all respondents.

- **Avoiding rushed or short data collection periods** – given the level of nonresponse bias climbs significantly if little flexibility is built into the data collection period, thus severely limiting a potential respondent's ability to answer, a survey collection period of at least two weeks will be ensured so that participants can choose to respond as per their schedule.

- **Send reminders to potential respondents** – reminder emails will be sent throughout the data collection period as this has been shown to effectively gather more completed responses. The first reminder email will be sent midway through the data collection period and the second near the end of the data collection period. Care will be taken to target only those who are yet to complete the survey.

- **Ensuring confidentiality** – it will be clearly expressed to respondents that the information they provide will be held securely; presented as part the whole sample and it will not be possible for it to be individually scrutinised.

In order to mitigate response bias, in this instance survey bias, four key actions will be considered: the way questions are worded, the selected question types and design, the structure of the survey, and its styling and colouring.

- **Question Wording** - all questions will be neutrally expressed. In addition, constructive peer review of the survey by Core Team members, MOPAN structures and members will provide an understanding of any potential problems with the questions that may make them confusing or erroneous for the assessment and target population.

- **Question Type and Design –** the selection of different forms of questions (rating scales, ranking, open-ended, closed-ended) and the options of answers provided for the respondent to choose from will be carefully considered to understand the strengths and weaknesses of each question type acknowledging that the selections made can have significant impact on the responses received.

- **Survey Structure** – review of survey structure will be undertaken at design stage in order to ensure a clear internal logic.

- **Survey styling** – styling is important to provide stimulus to the participant and avoid respondent fatigue. The use of colours and logos allows respondents to recognise a survey's legitimacy; a process of pretesting will be undertaken to ensure there are no issues with the choice of styling. Styling will be undertaken to ensure that the survey cannot be considered directed towards one demographic, looks neutral while still being inviting and professional.

Survey data analysis and ratings

The approach to analysing the survey data aligns with the rating scale, addressed in section 6. The initial thinking on essential aspects of the analysis and treatment of response data is outlined below.

First level data analysis

First level analysis of the survey data produces a basic summary of the sample data focusing on the frequency statistics and associated measures including the analysis of 'don't know'[17] responses and

[17] With the filters introduced to ensure knowledgeable respondents the level of 'don't know' will be reduced and such responses will be more likely to indicate that the respondent thinks the question is finely balanced (hence they will not make a judgement) rather than meaning they have insufficient knowledge to address the question.

missing data. This analysis also covers the qualitative content analysis of free-text responses to open questions.

Frequency calculation: Frequency tables are calculated for each micro-indicator using both weighted and unweighted response data. Missing data and 'don't know' responses are also tabulated. The table presents percentages for each frequency cell for rapid checking, ease of interpretation and use at the aggregation stages.

The frequency distribution information will be systematically treated to identify the balance of positive, negative and/or inconclusive responses.

Second level analysis

Second level analysis considers the difference in responses between sub-groups of the overall respondent data. This will be examined for differences in responses between respondent groups, but could also be carried out where the data suggests sub-groups of importance defined in other ways.

5.3 Interview and Consultations

Interviews and consultations comprise the third and fourth evidence lines for MOPAN 3.0, respectively. Given complementarities between them, they are addressed here in a single section of the Methodology Manual.

5.3.1 Interviews

Interviews need to be conducted systematically if the data gathers is to have maximum validity and make maximum contribution to the assessments.

The aims of the interviews are to:

- Deepen and interrogate the evidence from the document review
- Confirm or clarify trends/findings emerging from the document review
- Update on the multilateral organisations' on-going reform agenda and any key changes since documentation was analysed
- Provide contextual insight to clarify, refute and/or validate observations emerging from other lines of evidence/data sources (e.g. document review, survey data)
- Generate new evidence in areas where documentary and survey evidence is lacking
- Seek out explanations and factors (asking *why and how*)

Sampling

Interviews will be sought at two levels:

1) Headquarters level (during a one-week mission to HQ): some interviews will be conducted at senior levels, i.e. with senior management, heads of division/units, etc. The perspectives and insights of staff at the operational level are also important, however, since these often reveal valuable insights into whether and to what extent policy and strategic commitments are being implemented in practice. These insights will be gathered through consultations.

Box 8: Indicative interviewees – HQ

- Members of senior management
- Heads of key policy areas/units/divisions
- Heads of regional divisions
- Leads on internal reform processes
- Heads of key relevant business process (financial management, evaluation, performance management, human resources, programme design and approval, etc.)
- Donor liaison staff

2) Country/regional level: as the interviews will take place remotely, via phone or Skype, they are more likely to focus on mid- to senior level staff. Interviews will place a particular focus on confirming whether organisational commitments or reform processes are being implemented in reality, and whether these do, in fact, respond to country-level needs.

Box 9: Indicative interviewees – country/regional level

- Head of Office (Director, Representative or similar)
- Deputy Head of Office
- Senior Management
- Heads of Programme/policy areas
- Heads of key business processes (HR, finance, etc.)

Follow-up interviews may be required, to clarify any outstanding information gaps or to attempt to resolve any tensions or contradictions emerging at analysis stage. All interviewees will be asked whether or not they are prepared to undertake any follow-up interviews.

Methods for interviews

All interviews will apply a semi-structured framework, which covers the main areas for assessment adapted as appropriate, depending on a) tailoring per institution and b) findings from desk review and, if feasible, the survey. However, the structure will also permit flexibility, allowing new questions to be brought up during the interview as a result of what the interviewee says. Interview guides will be prepared in advance depending on individuals' functional responsibilities.

Interviews will apply standard ethical protocols,[18] with confidentiality and anonymity assured to interviewees, and without names being either attributed to data or in Annexes. Participants in group interviews (i.e. technical level staff) will be offered the opportunity of a follow up conversation, in case they wish to share information individually.

Data analysis and use

The information collected during the interviews will be analysed using a 'content analysis' approach, as for the documentary evidence. It will be treated as an independent line of evidence, but particularly as a source of insight, triangulation and verification, and a key source of evidence on the explanatory factors (why and how).

[18] E.g. the UNEG Ethical Guidelines

Interviews will not be transcribed in full, but data from them will be plotted into the composite analytical framework per relevant MI, to form the process of triangulation, above.

5.3.2 Consultations

Consultations are discrete from interviews, and focus on contextualisation/ triangulation/ validation/deepening. This distinction clarifies and makes explicit the different purposes and objectives of these two often blurred lines of evidence.

Approach

Within MOPAN 3.0, consultations are perceived as having greater scope for customisation and tailoring, with accordingly less structured tools applied than for interviews. The criteria for selecting respondents for consultation are also more flexible than those for interviews, being oriented mostly around gaps / shallow areas in data gathered.

Sampling

Consultations will take place at headquarter level only, and will be tailored to the needs of the individual assessment. They will focus on technical/administrative level staff, and will provide particularly valuable insights into the actual workings and culture of an organisation, and whether, how and to what extent policy and strategic commitments are being implemented in practice.

Consultations are applied with technical/administrative level staff, to explore a) contextual factors, b) verification of 'practice' (e.g., policy, strategy or business process implementation, etc.) and c) organisational culture.

Methodological approach

As a flexible tool, consultations will not prescribe a structure in advance, aiming rather for flexibility. Nonetheless, a clear question guide will be developed, linked to the relevant MI under exploration.

As for interviews, above, all consultations will be conducted to strict ethical standards, including protecting the anonymity and confidentiality of informants.

Data analysis and use

As for interviews, the data from consultations will not be transcribed, but relevant excerpts from it will be applied in the composite analytical framework as a means of deepening/ interrogation/ triangulation and verification.

SECTION 6: SCORING AND RATING SYSTEM

This section sets out the overall framework of the Scoring and Rating system, followed by its detailed approach and elements to be applied (which can be found in Annex C). The Scoring and Rating system will continue to be reviewed to ensure it is fit-for-purpose vis-a-vis MOPAN assessments.

Scoring and Rating is a critical dimension of the MOPAN assessment process, however, it is important to situate this within a wider narrative of the organisation's status at the current time. The wider text of the assessment report will therefore seek to provide a fuller illustration than can be provided by a numerical value alone.

Overall framework

The framework of the Scoring and Rating System for MOPAN 3.0 comprises the following:

1) Evidence from all four data streams (document review, survey, interviews and consultations) are brought together at analysis stage to comprise an aggregate evidence base per Micro Indicator (MI).

2) The scoring and rating system for each MI comprises a set number of **elements present**. When taken together, these elements demonstrate the presence or otherwise of **international best practice**.

3) The top end of the scoring and ratings scale reflects the implementation of the **full set of elements** which, when combined, represent **international best practice** against that MI (with explicit reference points identified for this where available). Conversely, the lower end of the scale (1) reflects **lesser presence** of these elements – and, consequently, weaker performance.

4) Added to this, to embed the commitment under MOPAN 3.0 to '**function over form**,' an organisation needs to have **actually implemented** the elements of best practice in order to score more highly.

5) Lower down the scale, an organisation either has **fewer** of the elements of international best practice in place, or has these in place (form) but not comprehensively implemented (function).

6) The MIs form the basis for a consolidated rating for each KPI, thus enabling a high-level indication of performance across the 12 KPI dimensions.

7) This approach allows us to situate organisations along the continuum of the '**performance journey**'; organisational maturity having emerged from the Interim Document Reviews thus far as a key theme.

Detailed approach:

The approach to Scoring and Rating under MOPAN 3.0 draws from the OECD Handbook on Constructing Composite Indicators: Methodology and User Guide (OECD 2008). This contains a set of technical guidelines aimed at helping constructors of composite indicators improve the quality of their indicators.

The approach to be adopted is as follows:

a) *Micro Indicator (MI) level – Within KPIs 1-8*

Each of the MOPAN 3.0 Micro Indicators (MIs) contain a number of elements agreed to represent 'international best practice. These vary in number.

The approach scores per element, on the basis of the extent to which an organisation implements the element, on a score of 1-4.

Thus:

Score per element	Descriptor
0	Element is not present
1	Element is present, but not implemented/implemented in zero cases
2	Element is partially implemented/implemented in some cases
3	Element is substantially implemented/implemented in the majority of cases
4	Element is fully implemented/implemented in all cases

Taking the average of the constituent scores per element, a rating is then calculated per MI. The ratings scale applied will be as follows:

Rating	Descriptor
3.01-4	Highly satisfactory
2.01-3	Satisfactory
1.01-2	Unsatisfactory
0-1	Highly unsatisfactory

b) *Micro Indicator (MI) level – Within KPIs 9-12*

KPIs 9-12 apply MIs drawn from the OECD DAC's Development Effectiveness Review model. These also apply a 0-4 numerical attribution system; with specific descriptors per score (see Annex D).

c) Aggregation to the KPI level

The same logic is pursued at aggregation to the KPI level, to ensure a consistent approach. Taking the average of the constituent scores per MI, a rating is then calculated per KPI.

The calculation for KPIs 1-8 is the same as for the MIs above, namely:

Rating	Descriptor
3.01-4	Highly satisfactory
2.01-3	Satisfactory
1.01-2	Unsatisfactory
0-1	Highly unsatisfactory

The ratings scale applied for KPIs 9-12 applies the same thresholds as for KPIs 1-8, for consistency, but will pitch scores to the middle of the threshold value (to guard against skewing in favour of higher ratings).

Elements of 'good practice' for Micro Indicators

Annex D provides the elements of good practice agreed by MOPAN members as the basis for assessment against the Micro Indicators (MIs) for MOPAN 3.0 and the reference points applied. Some common reference points apply to many or most MIs. These are as follows:

1) Busan Partnership for Effective Development Cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)

2) The resolution on the Quadrennial Comprehensive Policy Review (A/RES/67/226)

3) UNEG Norms and Standards for Evaluation, 2016

4) OECD-DAC Evaluating Development Cooperation, Summary of Key Norms and Standards, Second Edition

5) OIOS Inspection and Evaluation Manual, 2014

6) UNDG Results Based Management Handbook, 2011

7) The descriptors applied under the former Common Approach

Where no reference point is readily available, information from the first cycle of MOPAN 3.0 assessments has been applied (namely from the Document Reviews), alongside a sample of relevant items produced by agencies and, as appropriate, professional knowledge.

SECTION 7: TRIANGULATION AND VALIDATION

MOPAN 3.0 will apply multiple evidence sources per indicator. Section 6 described the use of multiple evidence sources which, where available, will then generate a composite rating.

Ratings will need to be supported by robust analytical strategies, to ensure validity and reliability, and therefore credibility of the evidence base. The following strategies will therefore be applied:

Triangulation

Triangulation is the process of using multiple data sources, data collection methods, and/or theories to validate research findings.[19] By using more than a single approach to address a research question, the risk of bias is reduced, and the chances of detracting errors or anomalies increased.[20]

Denzin (1978) and Patton (2001) identify four types of triangulation: methods triangulation, triangulation of sources, analyst triangulation and theory/perspective triangulation.[21] Of these, MOPAN 3.0 will apply three approaches, as follows:

1) ***Methods triangulation*** will occur by checking the consistency of findings generated by different data collection methods. This may occur by deepening enquiry via the sequential use of different data sources (for example, exploring findings from the document review through survey/interview and consultation data). Second, by interrogating data where diverging results arise (e.g. the document review shows evidence of a particular policy in place and being used, but survey data indicates little knowledge or use of the policy).

2) ***Triangulation of sources*** will occur by examining the consistency of different data sources within the same method. For example, by comparing through document review the application of corporate policies at country or regional level.

3) ***Analyst triangulation*** calls for the use of multiple analysts to review findings for the document review and composite analytical phases. This will allow for a consistent approach to interpretive analysis.

The Evidence Density Matrix, available in Annex C, sets out the anticipated evidence sources per MI, to be adapted per assessment. The approach adopted is that at least one primary source of evidence is allocated per MI. Where this is not a documentary source, at least two other forms of primary evidence must be available, i.e. survey/interview/consultation.

---

[19] The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment. Note: by combining multiple data-sources, methods, analyses or theories, evaluators hope to overcome the bias that comes from single informants, single methods, single observer or single theory studies.' OECD DAC (2002) Glossary of Terms on Results Based Management

[20] See for example: Morra-Imas, L and Rist, R (2009), The Road to Results: Designing and Conducting Effective Development Evaluations Washington DC: World Bank Publications. Also International Program for Development Evaluation Training (IPDET) Building Skills to Evaluate Development Interventions (undated) The World Bank Group, Carleton University and IOB/Ministry of Foreign Affairs, the Netherlands

[21] Denzin (1978) and Patton (2001) identify four types of triangulation – Methods triangulation; triangulation of sources; analyst triangulation and theory/perspective triangulation. Denzin, N (1978) Sociological Methods New York: McGraw Hill. Patton, M. Q. (2001) Qualitative Evaluation and Research Methods (2nd Edition). Thousand Oaks, California: Sage.

To aid comprehension of how analysis will take place, the following anonymised example is provided of an ‘evidence box’ (Box 10) at the end of the assessment process, with scoring applied for the MI.

Box 10: Sample evidence box

MI 1.1 Strategic plan and intended results based on a clear long term vision and analysis of comparative advantage			
Element	Score	Narrative	Source Documents
Element 1: The Strategic Plan (or equivalent) contains a long term vision	4	The multilateral organisation's current Strategic Plan (2014-2017) outlines a clear, long term vision and intended results. It identifies the organisation's overarching aim as…. The accompanying results framework outlines a set of seven outcome areas (listed) to monitor progress. An independent study, conducted in 2013, found the Strategic Plan to be "coherent and technically sound." The Theory of Change includes a comprehensive analysis of the organisation's comparative advantage in achieving its overarching aims. Its specific comparative advantages are identified as (listed). The accompanying results framework operationalises the Strategic Plan. It applies outputs to each of the outcome areas, and includes indicators and targets to monitor progress. Some gaps in logic were highlighted by the 2013 independent study, specifically a lack of clarity regarding the means through which the seven outcome areas converge to achieve impact. Additionally, the study found that in practice, and given the decentralised nature of the organisation, it is challenging to identify how the Strategic Plan will serve to shape decisions at the country level. Adjustments, (listed) have been subsequently implemented to address these shortcomings have since been introduced. Interviews and consultations found the Strategic Plan to be widely known and owned across the organisation, with staff able to clearly articulate their own roles in relation to the realisation of its objectives, and their contributions to specific corporate results. Survey data found that 113/130 responses who answered this question agreed that the Strategic Plan articulated a long-term vision for the organisation, and 96/127 relevant responses stated that it articulated clear intended results which were geared to the long term vision. The Strategic Plan was reviewed in 2016, midway through its implementation. A final evaluation of it is planned to start in early 2017.	(listed)
Element 2: The vision is based on a clear analysis and articulation of comparative advantage	4		
Element 3: The Strategic Plan operationalizes the vision, including defining intended results	3		
Element 4: The Strategic Plan is reviewed regularly to ensure continued relevance	4		
Overall Score:	3.75		
Overall Rating:	Highly satisfactory		High confidence

Validation

Validation of findings will occur at several points, as the Learning and Engagement strategy included in Section 12 will explain. Methods will include:

- The use of external evaluations and assessments of the organisations to help validate or question the findings on performance areas
- Debrief to assessed organisations/MOPAN members at relevant points, to hear the resonance of findings and judgements with insider knowledge of the organisation, and to test the accuracy of findings
- Validation of findings within the MOPAN network and revision of draft reports taking into account feedback from members
- The sharing of reports with multilateral organisations, and the taking of account of comments and feedback into consideration for the final draft of the report

SECTION 8: EVIDENCE MANAGEMENT

Through transparent management of evidence, the MOPAN 3.0 methodology aims to contribute to the credibility of MOPAN 3.0 assessments by making explicit and transparent the evidence base on which judgements are reached, including the signalling of any gaps

Evidence generated through MOPAN assessments will have three main forms:

1) the library of documentation provided for analysis

2) completed evidence box per MI (see Box 10 above)

3) a composite analytical Annex to the main report, which brings together the individual 'evidence boxes' and provides a narrative of evidence per KPI.

These forms of evidence will be made available to MOPAN members through the Repository, and also through the assessment reports themselves, which will include all three items (a Bibliography, and a full Annex containing points 2) and 3), referenced above).

SECTION 9: STRENGTHS AND WEAKNESSES OF THE METHODOLOGY

Building on the learning from Cycle 1 of the assessments, the expected strengths and limitations of the MOPAN 3.0 methodology are as follows. They will be confirmed and added to as Cycle 2 takes place. Indicatively, these are as follows:

Overall Strengths

- Applies clear and explicit operating principles, of credible, fair and accurate assessments – with clear strategies to achieve these
- Underpinned by solid theoretical foundations and a theory of change which, while supported by empirical evidence, is continuously tested and refined
- In tune with the current context in terms of the strategic discourse, policy emphases, reform processes and global development agenda
- Embedding key principles, criteria and commitment in terms of aid and development effectiveness including the commitments of the Busan Partnership and e.g. the International Humanitarian Principles as well as international criteria for humanitarian and development assistance
- Building on available frameworks for bilateral assessments (thereby with the intent of reducing the need for independent bilateral assessments)
- Applying multiple sources of data (stakeholder surveys, document review, interviews) against a single indicator, to ensure greater robustness of analysis and ratings
- Retaining the emphasis on stakeholder experience, in order that MOPAN members can better understand how the workings of multilateral organisations are understood at field and HQ level (also in line with principles of mutual accountability)
- Places an emphasis on engagement and learning, through the development and application of a separate Learning and Engagement Strategy (see Section 12)
- Places an emphasis on transparency, through the development and application of a full Evidence Management Strategy
- Has sound strategies for triangulation and validation, including engagement with the multilateral organisation and MOPAN members to confirm the resonance and the accuracy of the report
- Applies independent evaluative information for the assessment of results (therefore embedding the substance of the OECD DAC's Development Effectiveness Review)
- Places an emphasis on the quality, as well as the quantity, of the results generated by multilateral organisations, through seeking out information on their relevance, inclusiveness, sustainability and the cost-efficiency of their achievement

Overall Limitations

- The methodology, whilst it has taken into account members' own bilateral assessments, cannot always adequately cover individual members' specific needs e.g. on alignment with national priorities on in particular, member-specific areas. Therefore, it cannot replace all individual bilateral assessment tools (some of which are in any case required by e.g. Ministerial directives).
- The Common Approach framework was initially designed primarily, although not exclusively, to assess multilateral organisations with operational country programmes. MOPAN 3.0 has been designed with a wider focus, in order that it can encompass organisations with a largely normative or centralised operating model, and take account of regionally-focused structures and programming also. Nonetheless, careful checking will be required to ensure that indicators suitably encompass organisations with different mandates and operating models.
- The countries selected for review each year are based on established MOPAN criteria but comprise only a small proportion of each institution's operations, thus limiting broader generalisations. MOPAN 3.0 will be explicit in its reporting here.

- The highly constricted timeframe for MOPAN 3.0, combined with the upscaling of the number of organisations to be assessed, means that lines of evidence (e.g. document review) are of a necessarily smaller scale than would be feasible in e.g. a full institutional evaluation. This has been mitigated as far as possible through a systematic design, and the use of broad teams to cover as much ground as feasible within the timescale, but assessments – are not, therefore, comprehensive.
- Building on the learning from Cycle 1 of MOPAN 3.0 assessments, there is a need to ensure triangulation with a range of external sources - including institutional reviews and evaluations conducted by other agencies; and other sources of triangulation where available.

Possible limitations (and risks) per method include:

Survey

- The completion of the survey is reliant on MOPAN members and the assessed organisations identifying informants. Whilst efforts will be sought to identify and secure responses from the most informed and knowledgeable individuals (through the principle of 'right questions to the right people), a sufficient population for external validity of informants with 'reasonably expected knowledge' of the organisation cannot be guaranteed ex-ante.
- All perception surveys inevitably carry risks of differential interpretation by respondents on questions and ratings scales. This is particularly the case where surveys are conducted on a cross-national basis.[22] To mitigate this, questions and rating scale definitions will be presented in 'Plain English' (clearly and concisely) and professionally/expertly translated where necessary, and piloted with a global test group for acceptable accuracy of interpretation.
- To mitigate the risk of 'central tendency bias', where respondents gravitate towards a 'middle ground' score, the survey response scale is constructed on a symmetric 1-4 'forced choice' basis with an additional option for 'do not know/cannot respond'.
- Recognising availability of respondents is critical to the completion of the survey. Where possible and feasible, surveys will not be deployed at a time that is 'known' to be a quiet period (international holidays, etc.). The deployment date and timeframe for completion will be agreed with MOPAN members and the assessed organisations.

Document review

- Time constraints will necessarily mean that the sample of documentation which can be reviewed will be strategic/limited rather than comprehensive. Efforts will be made through other evidence lines (interviews , consultations, surveys) to address any gaps, but agreement on prioritisation will need to take place between the assessed organisation and the assessment team, via the Institutional Lead/ Secretariat.
- The document review component works within the confines of an organisation's disclosure policy. Where there are gaps due to unavailability of organisational documents, this will be explicitly reported.
- Evaluations included will have been generated (i.e. approved/ published) in the last two years (2015-2016) in order to apply as up to date information as possible. However, given finalization and

[22] See for example Harzing, A.W.; Reiche B.S.; Pudelko, M. (2012) Challenges in International Survey Research: A review with illustrations and suggested solutions for best practice, *European Journal of International Management*, vol. 5, no. 4, pp.

approval processes, there is necessarily a time lag between the results information available from evaluations, and those generated by the organisation in the recent period. To mitigate this, management performance information will also be applied to the development/humanitarian effectiveness component of the indicator framework (KPIs 9-12), but this will be explicitly flagged and treated as such, including the use of the hierarchy/triangulation strategies above, and incorporating external reviews and evaluations where available.

- Documentary evidence may well not capture reform processes currently underway, or how these are filtering down to country/regional level. Here, other evidence lines, such as the survey and interviews/consultations will be required to address the gap, and due acknowledgement – even if not robust documentary evidence – will need to be provided in reporting.

SECTION 10: MULTILATERAL ORGANISATION AND COUNTRY SELECTION

Multilateral Organisation Selection

MOPAN selects multilateral organisations for assessment on the basis of the following criteria:

- Perceived importance and interest to all MOPAN members
- Medium-term strategic planning (or equivalent) and replenishment cycles – with a view to assessing organisations prior to the planning process or the start of a replenishment negotiation process
- A mix of international financial institutions (IFIs), UN funds, programmes, specialised agencies, and humanitarian organisations

Their selection follows a dual-track process of a) the application of these criteria by Members through a process orchestrated by the MOPAN Secretariat; and b) a sampling process, based on clear criteria, conducted by IOD PARC as part of its inception work.

Country Selection

MOPAN 3.0 seeks findings from a number of countries (roughly around 12-16 per assessment cycle, however the number per specific organisation will vary depending on operations). The methodology explicitly does not seek external generalisability through country selection. It recognises that the limited selection of countries with which it is feasible to engage during the assessment process prevents generalisability in either statistical or theory-based terms. Nonetheless, a focus on the country level is valid, and important, for a range of reasons:

i. Triangulating and verifying information generated at corporate level
ii. Moving beyond 'form' to assess whether corporate policies, strategies and systems, rather than simply being present, are taken up and applied at country level
iii. Deepening enquiry on e.g. dimensions of comparative advantage for multilateral organisations, particularly in contexts where these dimensions come to the fore, such as in conflict-affected or fragile situations, or in challenging governance contexts.

For MOPAN 3.0, the existing criteria for country selection (as applied under previous assessment rounds) have been adapted below in Table 5.

Table 5: Application of Selection Criteria (countries)

1. Multilateral organisation presence in-country	Retained, and interpreted in the sense of 'scale and type of multilateral organisations' programming and operations in the country' since field visits are not envisaged under the MOPAN 3.0 methodology (unless data proves inaccessible by any other means)
2. Presence and availability of MOPAN members;	Discarded, for two reasons: • Presence or absence of members does not constitute a methodologically sound basis for country selection • Members may well be less 'present' in some contexts such as middle income countries, or fragile/conflict-affected situations, or in challenging governance environments. Yet these country contexts are arguably some of the most relevant environments for exploring the roles and comparative advantages of

	multilateral agencies, as above – and consequently their organisational and development effectiveness.
3. No recent inclusion in the survey;	Retained
4. The need for geographical spread;	Retained
5. Mix of low income and middle income countries (middle income countries being subdivided into lower middle and upper middle).	Retained but expanded to reflect an increased number of parameters for diversity (income status being only one axis along which the comparative advantage of multilateral organisations can be seen). These are: a. Gross National Income per Capita (World Bank 2014) b. Aid flows (OECD DAC 2014) c. Human Development Index score (UNDP 2014) d. Inequality - GINI coefficient (World Bank 2014) e. Gender and Development Index scores (UNDP 2014) f. Harmonised List of Fragile Situations (World Bank 2015) g. Political Stability and Absence of Violence/Terrorism (World Bank Governance Indicator) h. Rule of Law (World Bank Governance Indicator) i. Control of Corruption (World Bank Governance Indicator) j. Role as donor and/or recipient of ODA (OECD DAC AidStats)
6. *Additional* countries undertaking mutual assessment reviews of progress in implementing commitments[23]	Retained

The adopted parameters were then applied in sequence, as follows in Table 6 below.

Table 6: Parameters applied for country selection

Stage	Parameters applied
1	Presence of the multilateral organisations selected for assessment under Cycle 1 in countries classified as 'developing' by the OECD grouped by region (Africa, Americas, Australasia, Middle East and Europe).
2	Exclusion of countries having participated in the survey in 2014
3	Second-tier indicators as follows: a. GNI b. HDI c. GINI d. GDI e. Fragile Situation f. Governance Indicators

[23] As defined and monitored by the Global Partnership for Effective Development Co-operation. Such mutual assessments can provide complementary information for MOPAN about the performance of multilateral organisations at the country level.

	g.	Donor/recipient status
	h.	Inclusion in mutual assessment reviews

Specifically, the first stage of the process entailed a mapping exercise in which the *presence* of each of the confirmed multilateral organisations undergoing assessment in all the countries classified as 'developing' by the OECD. The purpose of the exercise was to ascertain 'multilateral organisation density' by country, through determining which multilateral agencies are present in which countries, as well as *how many* are present in each country.

Multilateral organisation density was used as the primary indicator for developing the first cut of the sample in which countries with low levels of multilateral organisation densities were excluded from the sample. Adequate representation across geographical regions was achieved through adjusting downwards the cut-off point at which countries were eliminated for regions with lower levels of multilateral organisation density. For example, Africa's comparatively high levels of overall multilateral organisation density meant a correspondingly high cut-off point at 10; any African country with less than 10 multilateral organisations present was therefore eliminated. On the other hand, in the Middle East, where multilateral organisation density was considerably lower, the cut-off point was adjusted to 8, in order to ensure adequate representation of Middle Eastern countries in the sample.

The next stage of the sampling process entailed further criteria-based elimination of countries in order to produce a 'long list' of countries deemed most relevant for assessment. This necessitated the collection of data pertaining to seven indicators for each country; GNI, HDI score, GINI coefficient, GDI, fragile situation (World Bank harmonised list), World Bank governance indicators (rule of law, political stability and control of corruption) and donor/recipient status.

These indicators are then systematically scrutinised for each country to develop the preliminary long and short lists (ranging from approx. 35 to 10, respectively), which are considered most relevant for country assessments. This is achieved by stratifying the countries according to the above indicators, and prioritising indicators in the following order:

i. GNI: all countries classified as 'high income' were eliminated, approximately equal representation of low, middle and upper middle income countries is the intention
ii. HDI, GINI and GDI: further elimination was guided by the need to ensure a balanced mix of these indicators. For countries classified as middle income, precedence is given to those displaying higher levels of inequality as measured by the GINI coefficient
iii. Fragile situation: the inclusion of several countries classified as existing in 'fragile situations' by the World Bank is ensured
iv. Governance indicators: efforts are made to ensure as much representation across each category as possible, in a bid to capture a variety of operating environments relevant to multilateral cooperation

It is important to include countries demonstrating 'contexts of interest' in relation to multilateral cooperation, and specifically, those where the potential comparative advantages of multilateral organisation, as outlined in Section 3, could reasonably be expected to come to the fore. 'Contexts of interest' include:

i. Fragile situations

ii. Countries with particularly low scoring governance indicators
iii. Countries experiencing or having recently experienced significant humanitarian crisis

MOPAN members agree and confirm the final list of countries per assessment cycle.

SECTION 11: REPORTING

Key principles of the report seek to ensure that it is accessible, concise and clear, and provides a contextualised rating.

Individual institutional reports are produced for each multilateral organisation assessed. These comprise a headline report (up to 45 pages) summarising the main narrative of the assessment, specifically reporting the key findings and conclusions against the 12 KPIs. A full set of Annexes contain the detailed findings against each MI.

The reports do not contain 'country chapters'. Instead, data acquired from country level through the document review and survey process, plus from any relevant interviews, is applied across analysis against each MI and KPI.

Key features of final assessment reports are:

- Accessible (minimising the use of jargon and ensuring accessible presentation of evidence)
- Clear (a clear and crystallised narrative for the main organisational report, with supporting evidence available elsewhere)
- Contextualised rating, which does not simply provide a number in isolation, but which grounds this in an evidence-based narrative of organisational change

SECTION 12: LEARNING AND ENGAGEMENT

A key component of the mission of MOPAN, as currently articulated, is to "contribute to organisational learning within and among multilateral organisations, their direct clients/partners and other stakeholders." Yet, a conclusion of the 2013 evaluation was that the usefulness of MOPAN's assessments to the multilateral organisations is limited. To address this issue, two initiatives are taken: 1) a formal communication strategy (Secretariat) and 2) a learning/engagement agenda (Technical Working Group and external service providers). These strategies will be refined and adjusted as the assessment process and cycles unfold and more experience is gained, and responsibilities will continue to be reviewed and refined.

In order to ensure that the assessment process and product spur organisational reflection and change, consulting key stakeholders, including MOPAN members and multilateral organisations, to ensure discussion, agreement and endorsement of key issues, findings and products at key stages both prior to and during the assessment cycle is critical. The current engagement/organisational learning strategy complements the communication strategy developed and implemented by the Secretariat; and more broadly defines and highlights opportunities to maximise organisational learning and clarify the responsibility of the respective stakeholders.

Opportunities for learning in the design of MOPAN 3.0

The design of MOPAN 3.0 presents a range of opportunities for learning. This may occur at both the theoretical and the empirical level on different dimensions, including:

- The dynamics and characteristics of the multilateral system, such as:
    - Characteristics and specificities of individual organisations
    - Characteristics of MOPAN and its members
- Methodological/evaluative issues in assessments, including the use of case studies
- In-process dialogue on various aspects of the assessments and the findings
- Best practices and benchmarking highlighted throughout the assessments

Opportunities are strengthened through the MOPAN 3.0 approach, which is sequenced and designed to include regular reflective feedback, on different issues with different groups of people. Each layer of information in the sequence of the assessment process will inform the next, as far as feasibility permits. So the survey, for example, will be informed as much as possible from the document review and, if possible, survey. Each of these processes will contain a learning element, both in terms of substance and in terms of methodology.

Mutual learning may thus occur and insights be generated at different points in the process, described below; the aim is to seize opportunities for learning as early as possible.

Points of engagement with the membership

Engagement between the external service providers and various elements of the MOPAN structure and assessment are outlined in table 7 below.

Table 7: Engagement between MOPAN actors and external service providers

MOPAN structure	Focus of the collaboration	Expected Frequency of interaction
Secretariat	Dialogue and feedback on all key deliverables and processes	Intensive throughout
Technical Working Group	Methodology design and development; learning	Key points during the Inception phase and key reflective points on methodology during and following implementation
MOPAN Institutional Leads	Preparing for, and organizing, each assessment	Key points during the implementation phase
MOPAN Steering Committee	Participation in Steering Committee meetings for information and learning	As deemed necessary and appropriate

Embedding learning in the assessment process

There are lessons from evaluation that are worth keeping in mind for the MOPAN assessment process, including:

1) The World Bank documented findings on how learning is more likely to come from engagement rather than through the reading of documents,[24] and therefore applying as much focus on the process as on the final product is integral.

2) Lessons and insights are generated throughout the process as evaluators ask questions, probe issues, and present emerging findings for discussion with partners and stakeholders. Michael Quinn Patton argues that research on evaluation demonstrates that, “intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings [and that] they are more likely to understand and feel ownership if they’ve been actively involved. By actively involving primary intended users, the evaluator is preparing the groundwork for use.”[25]

3) Academic and management literature in the current era operates with two aspects of learning: the creation of new knowledge (exploration) and the use of existing knowledge from various sources (exploitation).[26] Organisations need to succeed in both exploration and exploitation and keep an

[24] Learning and Results in World Bank Operations: How the Bank Learns, July 2014.
[25] Patton, M.Q., Utilization Focused Evaluations, 2008.
[26] March, J.G.: Exploration and Exploitation in Organisational Learning, Stanford, 1991.

appropriate balance between them. This is an insight that is both valuable in the assessment of the organisations' knowledge and performance management, but also in MOPAN's own work, as the multiple lines of evidence and the preparation of various assessment tools and analytical frameworks provide opportunities for both types of learning, including through discussion of past practices, and reviewing and drawing on new global best practices.

In terms of engagement with the organisations, they have in the past been given several opportunities to provide inputs into the assessment process, including help in adapting the methodology to fit the mandates of each particular organisation, identifying appropriate survey respondents and providing feedback during the drafting stages of the Institutional Reports. This approach is continued as it helps raise awareness of the issues examined by MOPAN, and provides opportunities for dialogue and mutual learning, both in terms of methodological issues, and substantively, in terms of the findings, and the implications of these for the individual multilateral organisation.

An indicative table of engagement with the multilateral organisations is provided in Table 8 below.

Table 8: Indicative engagement with multilateral organisations

Stage of the Assessment cycle	Nature of the engagement	Ways of facilitating the learning element
Notification & Kick Off	Inception mission and necessary calls	Sharing the thinking behind the learning opportunity of MOPAN 3.0; information sharing from the side of the organisation
Preparation & Document Review	Introductory Briefing Meeting at Organisation HQ Customisation of the assessment framework	The Briefing meeting will ideally take place aligned to another event of the organisation thereby facilitating wider awareness and interest on the MOPAN assessment. Introduce the generic Theory of Change as an important frame for the learning.
Survey, Interviews and Consultations	Customisation of the survey and interview and consultation schedule Tailored consultations	Debrief with the organisation on emerging themes at the end of the HQ visit
Report Writing	-	-
Commentary on Draft Report	Opportunity to comment on draft report	Discussion around key findings of the report
Presentation & Dissemination	Presentation of the institutional report and findings to senior management	Dissemination of a summary of the key points and outcomes of the discussion with the organisation

Sharing learning through reports

Notwithstanding the learning that is inherent in the process itself, the final published reports also play a key role both in terms of formalising the findings and providing an opportunity for each organisation to respond, and in terms of sharing the report with various stakeholders, who have not been directly involved in the process.

The presentation of the final report(s), and respective dialogue with management and governance bodies in the assessed organisations is the role of MOPAN members, with the external service providers backstopping in a supporting, technical role.

Other written documents that embody or strengthen learning include this Methodology Manual, a methodological lessons learned paper, as well as ‘intermediary products’ that will, as of time of publication, not be for the general public, such as the desk review reports, survey results and various templates.

Final reports are developed with the intention of being accessible and reader-friendly, and will have well evidenced findings that highlight both areas for improvement and areas of high performance and good practice in each multilateral organisation.

Annex A: Indicator Framework

PERFORMANCE AREA: **STRATEGIC MANAGEMENT**

Clear strategic direction geared to key functions, intended results and integration of relevant cross-cutting priorities

KPI 1: Organisational architecture[27] and financial framework enable mandate implementation and achievement of expected results

1.1 Strategic plan and intended results based on a clear long term vision and analysis of comparative advantage

1.2 Organisational architecture congruent with a clear long term vision and associated operating model

1.3 Strategic plan supports the implementation of wider normative frameworks and associated results, (e.g. the quadrennial comprehensive policy review (QCPR), replenishment commitments, or other resource and results reviews)

1.4 Financial framework (e.g. division between core and non-core resources) supports mandate implementation

KPI 2: Structures and mechanisms in place and applied to support the implementation of global frameworks for cross-cutting issues at all levels

2.1 Corporate/sectoral and country strategies respond to and/or reflect the intended results of normative frameworks for cross-cutting issues

2.1a Gender equality and the empowerment of women

2.1b Environmental sustainability and climate change

2.1c Good governance (peaceful and inclusive societies for sustainable development, reduced inequality, provide access to justice for all and build effective, accountable and inclusive institutions at all levels)

2.1d Human Rights

2.1e Any other cross-cutting issue included in organisational mandates/commitments

[27] *Organisational Architecture is "a theory of the firm, or multiple firms, which integrates the human activities and capital resource utilization within a structure of task allocation and coordination to achieve desired outcomes and performance for both the short run and the strategic long run" (Burton and Obel, 2011a, 2011b).*

PERFORMANCE AREA: **OPERATIONAL MANAGEMENT**

Assets and capacities organised behind strategic direction and intended results, to ensure relevance, agility and accountability

KPI 3: Operating model and human/financial resources support relevance and agility

3.1 Organisational structures and staffing ensure that human and financial resources are continuously aligned and adjusted to key functions

3.2 Resource mobilization efforts consistent with the core mandate and strategic priorities

3.3 Aid reallocation / programming decisions responsive to need can be made at a decentralised level

3.4 HR systems and policies performance based and geared to the achievement of results

KPI 4: Organisational systems are cost- and value-conscious and enable financial transparency/accountability

4.1 Transparent decision-making for resource allocation, consistent with strategic priorities

4.2 Allocated resources disbursed as planned

4.3 Principles of results based budgeting applied

4.4 External audit or other external reviews certifies the meeting of international standards at all levels, including with respect to internal audit

4.5 Issues or concerns raised by internal control mechanisms (operational and financial risk management, internal audit, safeguards etc.) adequately addressed

4.6 Policies and procedures effectively prevent, detect, investigate and sanction cases of fraud, corruption and other financial Irregularities

PERFORMANCE AREA: **RELATIONSHIP MANAGEMENT**

Engaging in inclusive partnerships to support relevance, to leverage effective solutions and to maximise results (in line with Busan Partnerships commitments)

KPI 5: Operational planning and intervention design tools support relevance and agility (within partnerships)

5.1 Interventions aligned with national/regional priorities and intended national/regional results

5.2 Contextual analysis (shared where possible) applied to shape the intervention designs and implementation.

5.3 Capacity analysis informs intervention design and implementation, and strategies to address any weakness found are employed

5.4 Detailed risk (strategic, political, reputational, operational) management strategies ensure the identification, mitigation, monitoring and reporting of risks

5.5 Intervention designs include the analysis of cross-cutting issues (as defined in KPI2)

5.6 Intervention designs include detailed and realistic measures to ensure sustainability (as defined in KPI 12)

5.7 Institutional procedures (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.) positively support speed of implementation

KPI 6: Works in coherent partnerships directed at leveraging and/or ensuring relevance and catalytic use of resources

6.1 Planning, programming and approval procedures enable agility in partnerships when conditions change

6.2 Partnerships based on an explicit statement of comparative advantage e.g. technical knowledge, convening power/partnerships, policy dialogue/advocacy

6.3 Clear adherence to the commitment in the Busan Partnership for Effective Development Cooperation on use of country systems

6.4 Strategies or designs identify synergies, to encourage leverage/catalytic use of resources and avoid fragmentation

6.5 Key business practices (planning, design, implementation, monitoring and reporting) co-ordinated with other relevant partners (donors, UN agencies, etc.), as appropriate

6.6 Key information (analysis, budgeting, management, results etc.) shared with strategic/implementation partners on an ongoing basis

6.7 Clear standards and procedures for accountability to beneficiaries implemented

6.8 Participation with national and other partners in mutual assessments of progress in implementing agreed commitments

6.9 Deployment of knowledge base to support programing adjustments, policy dialogue and/or advocacy

PERFORMANCE AREA: **PERFORMANCE MANAGEMENT**

Systems geared to managing and accounting for development and humanitarian results and the use of performance information, including evaluation and lesson-learning

KPI 7: Strong and transparent results focus, explicitly geared to function

7.1 Leadership ensures application of an organisation-wide RBM approach

7.2 Corporate strategies, including country strategies, based on a sound RBM focus and logic

7.3 Results targets set based on a sound evidence base and logic

7.4 Monitoring systems generate high quality and useful performance data

7.5 Performance data transparently applied in planning and decision-making.

KPI 8: Evidence-based planning and programming applied

8.1 A corporate independent evaluation function exists

8.2 Consistent, independent evaluation of results (coverage)

8.3 Systems applied to ensure the quality of evaluations

8.4 Mandatory demonstration of the evidence base to design new interventions

8.5 Poorly performing interventions proactively identified, tracked and addressed

8.6 Clear accountability system ensures responses and follow-up to and use of evaluation recommendations

8.7 Uptake of lessons learned and best practices from evaluations.

PERFORMANCE AREA: **RESULTS**

Achievement of relevant, inclusive and sustainable contributions to humanitarian and development results in an efficient way.

KPI 9: Achievement of development and humanitarian objectives and results e.g. at the institutional/corporate-wide level and regional/country level, with results contributing to normative and cross-cutting goals.

9.1 Interventions assessed as having achieved their stated development and/or humanitarian objectives and attain expected results

9.2 Interventions assessed as having realised the expected positive benefits for target group members

9.3 Interventions assessed as having contributed to significant changes in national development policies and programs (policy and capacity impacts), or needed system reforms

9.4 Interventions assessed as having helped improve gender equality and the empowerment of women

9.5 Interventions assessed as having helped improve environmental sustainability/helped tackle the effects of climate change

9.6 Interventions assessed as having helped improve good governance (as defined in 2.1.c)

9.7 Interventions assessed as having helped improve human rights

FACTORS[28] – REASONS WHY RESULTS WERE ACHIEVED OR NOT

- **External – context-related reasons**
  - *Operating context*
  - *Governance context*
  - *Financial context*
  - *Partner (national/regional partner, donor, wider multilateral) context*
- **Internal (signal positive or negative)**
  - *Policy issues;*
  - *Programme or project design;*
  - *Objectives /targets – appropriate, realistic;*
  - *Financial resource issues;*
  - *Human resource issues;*
  - *Implementation challenges;*
  - *Oversight/governance of the institution;*
  - *Risk management;*
  - *Communication and decision-making systems;*
  - *Use of innovation (specify).*

KPI 10: Relevance of interventions to the needs and priorities of partner countries and beneficiaries, and extent to which the organisation works towards results in areas within its mandate

10.1 Interventions assessed as having responded to the needs / priorities of target groups

10.2 Interventions assessed as having helped contribute to the realisation of national development goals and objectives

[28] These will be extracted as available from the evidence (particularly evaluations), with a view to informing findings against a range of MIs and for later collation, rather than to be assessed or rated as discrete data

10.3 Results assessed as having been delivered as part of a coherent response to an identified problem

KPI 11: Results delivered efficiently

11.1 Interventions assessed as resource/cost efficient

11.2 Implementation and results assessed as having been achieved on time (given the context, in the case of humanitarian programming)

KPI 12: Sustainability of results

12.1 Benefits assessed as continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to recovery, resilience eventually, to longer-term developmental results

12.2 Interventions assessed as having built sufficient institutional and/or community capacity for sustainability, or have been absorbed by government

12.3 Interventions assessed as having strengthened the enabling environment for development

Annex B: Evidence Density Matrix

P = primary source of evidence; S = secondary

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
Performance Area: Strategic management: **Clear strategic direction geared to key functions, intended results and integration of relevant cross-cutting priorities**				
KPI 1: Organisational architecture[29] and financial framework enable mandate implementation and achievement of expected results				
1.1 Strategic plan and intended results based on a clear long term vision and comparative advantage	P		S	S
1.2 Organisational architecture congruent with a clear long term vision and associated operating model	P			
1.3 Strategic plan supports the implementation of wider normative frameworks and associated results, (e.g. the quadrennial comprehensive policy review (QCPR), replenishment commitments, or other resource and results reviews)	P			
1.4 Financial framework (e.g. division between core and no-core resources) supports mandate implementation	P		S	S
KPI 2: Structures and mechanisms in place to support the implementation of global frameworks for cross-cutting issues at all levels				
2.1 Corporate/sectoral and country strategies respond to and/or reflect the intended results of normative frameworks for cross-cutting issues 2.1a Gender equality and the empowerment of women	P	S	P	S

29 Organisational Architecture is "a theory of the firm, or multiple firms, which integrates the human activities and capital resource utilization within a structure of task allocation and coordination to achieve desired outcomes and performance for both the short run and the strategic long run" (Burton and Obel, 2011a, 2011b).

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
2.1b Environmental sustainability and climate change 2.1c Good governance *(peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels)*[30] 2.1d Human Rights 2.1e Any other cross-cutting issue included in organisational mandates/commitments				
Performance Area: Operational Management **Assets and capacities organised behind strategic direction and intended results, to ensure relevance, agility and accountability**				
KPI 3: Operating model and human/financial resources support relevance and agility				
3.1 Organisational structures and staffing ensure that human and financial resources are continuously aligned and adjusted to key functions	P	P	P	S
3.2 Resource mobilization efforts consistent with the core mandate and strategic priorities	P		S	S
3.3 Aid reallocation / programming decisions responsive to need can be made at a decentralised level	S	P	P	S
3.4 HR systems and policies performance based and geared to the achievement of results	P	S		

[30] Definition to be finalised following a meeting of the Technical Working Group in November 2015

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
KPI 4: Organisational systems are cost- and value-conscious and enable financial transparency/accountability				
4.1 Transparent decision-making for resource allocation, consistent with strategic priorities	P	S	P	S
4.2 Allocated resources disbursed as planned	P		S	
4.3 Principles of results based budgeting applied	P		S	
4.4 External audit or other external reviews certifies the meeting of international standards at all levels, including with respect to internal audit	P			
4.5 Issues or concerns raised by internal control mechanisms (operational and financial risk management, internal audit, safeguards etc.) adequately addressed	P	S	S	S
4.6 Policies and procedures effectively prevent, detect, investigate and sanction cases of fraud, corruption and other financial irregularities	P		S	S
Performance Area: Relationship Management **Engaging in inclusive partnerships to support relevance, to leverage effective solutions and to maximise results (in line with Busan Partnerships commitments)**				
KPI 5: Operational planning and intervention design tools support relevance and agility (within partnerships)				
5.1 Interventions aligned with national/regional priorities and intended national/regional results	P		S	S

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
5.2 Contextual analysis (shared where possible) applied to shape the intervention designs and implementation.	P		S	S
5.3 Capacity analysis informs intervention design and implementation, and strategies to address any weakness found are employed	P		S	S
5.4 Detailed risk (strategic, political, reputational, operational) management strategies ensure the identification, mitigation, monitoring and reporting of risks	P		S	S
5.5 Intervention designs include the analysis of cross-cutting issues (as defined in KPI2)	P		S	S
5.6 Intervention designs include detailed and realistic measures to ensure sustainability (as defined in KPI12)	P		S	S
5.7 Institutional procedures (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.) positively support speed of implementation	P	P	S	S
KPI 6: Works in coherent partnerships directed at leveraging and /or ensuring relevance and catalytic use of resources				
6.1 Planning, programming and approval procedures enable agility in partnerships when conditions change	P	P	S	S
6.2 Partnerships based on an explicit statement of comparative advantage e.g. technical knowledge, convening power/partnerships, policy dialogue/advocacy		P	P	P

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
6.3 Clear adherence to the commitment in the Busan Partnership for Effective Development Cooperation on use of country systems	P	S	S	
6.4 Strategies or designs identify synergies, to encourage leverage/catalytic use of resources and avoid fragmentation	S	P	P	P
6.5 Key business practices (planning, design, implementation, monitoring and reporting) co-ordinated with other relevant partners (donors, UN agencies, etc.), as appropriate	S	P	P	P
6.6 Key information (analysis, budgeting, management, results etc.) shared with strategic/implementation partners on an ongoing basis	P	P	S	S
6.7 Clear standards and procedures for accountability to beneficiaries implemented	P	P	P	P
6.8 Participation with national and other partners in mutual assessments of progress in implementing agreed commitments	P	P	S	S
6.9 Deployment of knowledge base to support programming adjustments, policy dialogue and/or advocacy	P	P	S	S
Performance Area: Performance Management **Systems geared to managing and accounting for development and humanitarian results and the use of performance information, including evaluation and lesson-learning**				
KPI 7: Strong and transparent results focus, explicitly geared to function				
7.1 Leadership ensures application of an organisation-wide RBM approach	P		S	S

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
7.2 Corporate strategies, including country strategies, based on a sound RBM focus and logic	P		S	S
7.4 Results targets set based on a sound evidence base and logic	P		S	S
7.5 Monitoring systems generate high quality and useful performance data	P		S	S
7.6 Performance data transparently applied in planning and decision-making.	P		S	S
KPI 8: Evidence-based planning and programming applied				
8.1 A corporate independent evaluation function exists	P			
8.2 Consistent, independent evaluation of results (coverage)	P		S	S
8.3 Systems applied to ensure the quality of evaluations	P		S	S
8.4 Mandatory demonstration of the evidence base to design new interventions	P	S	S	S
8.5 Poorly performing interventions proactively identified, tracked and addressed	P	S	S	S
8.6 Clear accountability system ensures responses and follow up to and use of evaluation recommendation	P		S	S
8.7 Uptake of lessons learned and best practices from evaluations.	P		S	S

Performance Area: Results **Achievement of relevant, inclusive and sustainable contributions to humanitarian and development results in an efficient way.**	Doc. Review
KPI 9 Achievement of development and humanitarian objectives and results e.g. at the institutional/corporate-wide level and regional/country level, with results contributing to normative and cross-cutting goals	
9.1 Interventions/activities assessed as having achieved their stated development and/or humanitarian objectives and attain expected results	P
9.2 Interventions/activities assessed as having realized the expected positive benefits for target group members	P
9.3 Interventions/activities assessed as having contributed to significant changes in national development policies and programs (policy and capacity impacts), or needed system reforms	P
9.4 Interventions/activities assessed as having helped improve gender equality and the empowerment of women	P
9.5 Interventions/activities assessed as having helped improve environmental sustainability/helped tackle the effects of climate change	P
9.6 Interventions/activities assessed as having helped improve good governance (as defined in KPI2, MI1.c)	P
9.7 Interventions/activities assessed as having helped improve human rights	P
Factors – reasons why results were achieved or not[31] • External – context-related reasons ○ Operating context ○ Governance context ○ Financial context ○ Partner (national/regional, donor, wider multilateral) context • Internal (signal positive or negative) ○ Policy issues	P

[31] These will be extracted as available from the evidence (particularly evaluations), with a view to informing findings against a range of MIs and for later collation, rather than to be assessed or rated as discrete data

○ Programme or project design; ○ Objectives /targets – appropriate, realistic ○ Financial resource issues; ○ Human resource issues; ○ Implementation challenges; ○ Oversight/governance of the institution; ○ Risk management; ○ Communication and decision-making systems ○ Use of innovation (specify)	
KPI 10 Relevance of interventions to the needs and priorities of partner countries and beneficiaries, and extent to which the organisation works towards results in areas within its mandate	
10.1 Interventions/activities assessed as having responded to the needs / priorities of target groups	P
10.2 Interventions/activities assessed as having helped contribute to the realisation of national development goals and objectives	P
10.3 Results assessed as having been delivered as part of a coherent response to an identified problem	P
KPI 11 Results delivered efficiently	
11.1 Interventions / activities assessed as resource/cost efficient	P
11.2 Implementation and results assessed as having been achieved on time (given the context, in the case of humanitarian programming)	P
KPI 12: Sustainability of results	
12.1 Benefits assessed as continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to recovery, resilience eventually, to longer-term developmental results	P
12.2 Interventions/activities assessed as having built sufficient institutional and/or community capacity for sustainability, or have been absorbed by government	P

12.3 Interventions/activities assessed as having strengthened the enabling environment for development	P

Annex C: Scoring & Rating System

KPI 1: Organisational architecture[32] and financial framework enable mandate implementation and achievement of expected results		
Micro Indicator	**Elements**	**Reference point(s)**
1.1 Strategic plan and intended results based on a clear long term vision and analysis of comparative advantage	1. A publicly available Strategic Plan (or equivalent) contains a long term vision 2. The vision is based on a clear analysis and articulation of comparative advantage 3. A strategic plan operationalizes the vision, including defining intended results 4. The strategic plan is reviewed regularly to ensure continued relevance	*OIOS evaluation methodology* *Common approach*
1.2 Organisational architecture congruent with a clear long term vision and associated operating model	1. The organisational architecture is congruent with the strategic plan 2. The operating model supports implementation of the strategic plan 3. The operating model is reviewed regularly to ensure continued relevance 4. The operating model allows for strong cooperation across the organisation and with other agencies 5. The operating model clearly delineates responsibilities for results	*Burton and Obel, various works, organisational Development – Principles. Processes, Performance by Gary N. Mc Lean*
1.3 The strategic plan supports the	1. The strategic plan is aligned to wider normative frameworks and associated results, including Agenda 2030, and others, such as the QCPR and the Grand Bargain (where	*Common approach*

[32] Organisational Architecture is "a theory of the firm, or multiple firms, which integrates the human activities and capital resource utilization within a structure of task allocation and coordination to achieve desired outcomes and performance for both the short run and the strategic long run" (Burton and Obel, 2011a, 2011b).

implementation of wider normative frameworks and associated results, including Agenda 2030 and others where applicable (e.g. the quadrennial comprehensive policy review (QCPR), Grand Bargain, replenishment commitments or other resource and results reviews)	applicable) 2. The strategic plan includes clear results for normative frameworks, including Agenda 2030, and others, such as the QCPR and the Grand Bargain (where applicable) 3. A system to track normative results is in place for Agenda 2030, and any other relevant frameworks, such as the QCPR and the Grand Bargain (where applicable) 4. The organisation's accountability for achieving normative results, including those of Agenda 2030, and any other relevant frameworks, such as the SDGs and their targets and indicators, the QCPR and the Grand Bargain (where applicable), is clearly established 5. Progress on implementation on an aggregated level is published at least annually	
1.4 Financial framework (e.g. division between core and non-core resources) supports mandate implementation	1. Financial and budgetary planning ensures that all priority areas have adequate funding in the short term or are at least given clear priority in cases where funding is very limited 2. A single integrated budgetary framework ensures transparency 3. The financial framework is reviewed regularly by the governing bodies 4. Funding windows or other incentives in place to encourage donors to provide more flexible/un-earmarked funding at global and country levels 5. Policies/measures are in place to ensure that earmarked funds are targeted at priority areas	*OECD Multilateral reports*
KPI 2: Structures and mechanisms in place and applied to support the implementation of global frameworks for cross-cutting issues at all levels		
2.1a Gender equality and the empowerment of women	1. Dedicated policy statement on gender equality available and showing evidence of use 2. Gender equality indicators and targets fully integrated into the organisation's strategic plan and corporate objectives 3. Accountability systems (including corporate reporting and evaluation) reflect gender equality indicators and targets 4. Gender screening checklists or similar tools used for all new Interventions 5. Human and financial resources (exceeding benchmarks) are available to address gender issues 6. Capacity development of staff on gender is underway or has been conducted	*UNSWAP; Gender evaluations (e.g. Joint evaluation of Joint Gender Programmes in the UN system).* *Common approach*
2.1b Environmental	1. Dedicated policy statement on environmental sustainability and climate change available	*UNDP/UNEP: Mainstreaming*

sustainability and climate change	and showing evidence of use 2. Environmental sustainability /climate change indicators and targets are fully integrated into the organisation's strategic plan and corporate objectives 3. Accountability systems (including corporate reporting and evaluation) reflect environmental sustainability and climate change indicators and targets 4. Environmental screening checklists/ impact assessments used for all new Interventions 5. Human and financial resources are available to address environmental sustainability and climate change issues 6. Capacity development of staff of environmental and climate change issues is underway or has taken place	*Environment and Climate for Poverty Reduction and Sustainable Development; UNDG: Mainstreaming Environmental Sustainability in Country Analysis and the UNDAF: World Bank: Environmental and Social Framework (2nd Draft July 2015)* *Common approach*
2.1c Good governance	1. Dedicated policy statement on the principles of good governance and effective institutions available and showing evidence of use 2. Indicators and targets related to the principles of good governance and effective institutions are integrated into the organisation's strategic plan and corporate objectives 3. Accountability systems (including corporate reporting and evaluation) reflect the principles of good governance and effective institutions 4. New interventions are assessed for relevant governance/institutional effectiveness issues 5. Human and financial resources are available to address the principles of good governance and issues related to effective institutions 6. Capacity development of staff on the principles of good governance and effective institutions is underway or has taken place	*World Bank Governance assessment framework; UNDP Planning a Governance Assessment: a guide to approaches, costs and benefits; UNDP Discussion Paper Governance for Sustainable Development – Integrating Governance in the Post-2015 Development Framework* *Common approach*
2.1d Human Rights	1. Dedicated policy statement on human rights available and showing evidence of use 2. Human rights indicators and targets fully integrated into the organisation's strategic plan and corporate objectives 3. Accountability systems (including corporate reporting and evaluation) reflect human rights indicators and targets 4. Human rights screening checklists or similar tools used for all new interventions 5. Human rights and financial resources (exceeding benchmarks) are available to address human rights issues	*The Human Rights Based Approach to Development Cooperation; Integrating human rights in development and in the economic sphere (OHCHR); ICERD, ICCPR, ICESCR, CEDAW, CAT, CRC, ICMW, CPED, CRPD* *Common approach*

	6. Capacity development of staff on is underway or has been conducted	
KPI 3: Operating model and human/financial resources support relevance and agility		
3.1 Organisational structures and staffing ensure that human and financial resources are continuously aligned and adjusted to key functions	1. Staffing is aligned with, or being reorganized to, requirements set out in the current Strategic Plan 2. Resource allocations across functions are aligned to current organisational priorities and goals, as set out in the current Strategic Plan 3. Internal restructuring exercises have a clear purpose and intent, aligned to the priorities of the current Strategic Plan	*Interim Document Reviews*
3.2 Resource mobilization efforts consistent with the core mandate and strategic priorities	1. Resource mobilization strategy/case for support explicitly aligned to current strategic plan 2. Resource mobilization strategy/case for support reflects recognition of need to diversify the funding base, particularly in relation to the private sector 3. Resource mobilization strategy/case for support seeks multi-year funding within mandate and strategic priorities 4. Resource mobilization strategy/case for support prioritises the raising of domestic resources from partner countries/institutions, aligned to goals and objectives of the Strategic Plan/relevant country plan 5. Resource mobilization strategy/case for support contains clear targets, monitoring and reporting mechanisms geared to the Strategic Plan or equivalent	*Sample resource mobilization strategies (UNHCR, UNICEF, WFP)*
3.3 Aid reallocation / programming decisions responsive to need can be made at a decentralised level	1. An organisation-wide policy or guidelines exist which describe the delegation of decision-making authorities at different levels within the organisation 2. *(If the first criterion is met)* The policy/guidelines or other documents provide evidence of a sufficient level of decision making autonomy available at the country level (or other decentralized level as appropriate) regarding aid reallocation/programming 3. Evaluations or other reports contain evidence that reallocation/programming decisions have been made to positive effect at country or other local level, as appropriate 4. The organisation has made efforts to improve or sustain the delegation of decision-making on aid allocation/programming to the country or other relevant levels	*Common Approach (with adjustment), sample of UN agency Resource Mobilization Strategies, Interim Document Reviews*

3.4 HR systems and policies performance based and geared to the achievement of results	1. A system is in place which requires the performance assessment of all staff, including senior staff 2. There is evidence that the performance assessment system is systematically and implemented by the organisation across all staff and to the required frequency 3. The performance assessment system is clearly linked to organisational improvement, particularly the achievement of corporate objectives, and to demonstrate ability to work with other agencies 4. The performance assessment of staff is applied in decision making relating to promotion, incentives, rewards, sanctions, etc. 5. A clear process is in place to manage disagreement and complaints relating to staff performance assessments	*Common Approach (with adjustment), Interim Document Reviews*
KPI 4: Organisational systems are cost- and value-conscious and enable financial transparency/accountability		
4.1 Transparent decision-making for resource allocation, consistent with strategic priorities	1. An explicit organisational statement or policy exists which clearly defines criteria for allocating resources to partners 2. The criteria reflect targeting to the highest priority themes/countries/areas of intervention as set out in the current Strategic Plan 3. The organisational policy or statement is regularly reviewed and updated 4. The organisational statement or policy is publicly available	*Common Approach (with adjustment), sample of agency resource allocation documentation*
4.2 Allocated resources disbursed as planned	1. The institution sets clear targets for disbursement 2. Financial information indicates that planned disbursements were met within institutionally agreed margins 3. Clear explanations are available in relation to any variances 4. Variances relate to external factors rather than internal procedural blockages	*Common Approach (with adjustment)*
4.3 Principles of results based budgeting applied	1. The most recent organisational budget clearly aligns financial resources with strategic objectives/intended results of the current Strategic Plan 2. A budget document is available which provides clear costings for the achievement of	*Common Approach (with adjustment), sample of agency Results Based Budgeting statements*

	each management result 3. Systems are available and used to track costs from activity through to result (outcome) 4. There is evidence of improved costing of management and development results in budget documents reviewed over time (evidence of building a better system)	
4.4 External audit or other external reviews certifies the meeting of international standards at all levels, including with respect to internal audit	1. External audit conducted which complies with international standards 2. Most recent external audit confirms compliance with international standards across functions 3. Management response is available to external audit 4. Management response provides clear action plan for addressing any gaps or weaknesses identified by external audit 5. Internal audit functions meet international standards, including for independence 6. Internal audit reports are publicly available	*Common Approach (with adjustment), sample of agency external audits*
4.5 Issues or concerns raised by internal control mechanisms (operational and financial risk management, internal audit, safeguards etc.) adequately addressed	1. A clear policy or organisational statement exists on how any issues identified through internal control mechanisms will be addressed 2. Management guidelines or rules provide clear guidance on the procedures for addressing any identified issues, including timelines 3. Clear guidelines are available for staff on reporting any issues identified 4. A tracking system is available which records responses and actions taken to address any identified issues 5. Governing Body or management documents indicate that relevant procedures have been followed/action taken in response to identified issues, including recommendations from audits (internal and external) 6. Timelines for taking action follow guidelines/ensure the addressing of the issue within twelve months following its reporting.	*Common Approach (with adjustment), Interim Document Reviews*
4.6 Policies and procedures effectively prevent, detect, investigate and sanction cases of fraud, corruption and other financial irregularities	1. A clear policy/guidelines on fraud, corruption and any other financial irregularities is available and made public 2. The policy/guidelines clearly define the roles of management and staff in implementing/complying with the guidelines 3. Staff training/awareness-raising has been conducted in relation to the policy/guidelines 4. There is evidence of policy/guidelines implementation, e.g. through regular monitoring	*Common Approach (with adjustment), Interim Document Reviews*

	and reporting to the Governing Body 5. There are channels/mechanisms in place for reporting suspicion of misuse of funds (e.g. anonymous reporting channels and "whistle-blower" protection policy 6. Annual reporting on cases of fraud, corruption and other irregularities, including actions taken, ensures that they are made public	
KPI 5: Operational planning and intervention design tools support relevance and agility (within partnerships)		
5.1 Interventions aligned with national/regional priorities and intended national/regional results	1. Reviewed country or regional strategies make reference to national/regional strategies or objectives 2. Reviewed country strategies or regional strategies link the results statements to national or regional goals 3. Structures and incentives in place for technical staff that allow investment of time and effort in alignment process	*Busan Partnership for effective development cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)* *Quadrennial Comprehensive Policy Review resolution*
5.2 Contextual analysis (shared where possible) applied to shape the intervention designs and implementation	1. Intervention designs contain a clear statement that positions the intervention within the operating context 2. Context statement has been developed jointly with partners 3. Context analysis contains reference to gender issues, where relevant 4. Context analysis contains reference to environmental sustainability and climate change issues, where relevant 5. Context analysis contains reference to governance issues, including conflict and fragility, where relevant 6. Evidence of reflection points with partner(s) that take note of any significant changes in context	*Common approach/Interim Document Reviews*
5.3 Capacity analysis informs intervention design and implementation, and strategies to address any weakness found are	1. Intervention designs contain a clear statement of capacities of key national implementing partners 2. Capacity analysis considers resources, strategy, culture, staff, systems and processes, structure and performance 3. Capacity analysis statement has been developed jointly where feasible	*Interim Document Reviews*

employed	4. Capacity analysis statement includes clear strategies for addressing any weaknesses, with a view to sustainability 5. Evidence of regular and resourced reflection points with partner(s) that take note of any significant changes in the wider institutional setting that affect capacity	
5.4 Detailed risk (strategic, political, reputational, operational) management strategies ensure the identification, mitigation, monitoring and reporting of risks	1. Intervention designs include detailed analysis of and mitigation strategies for operational risk 2. Intervention designs include detailed analysis of and mitigation strategies for strategic risk 3. Intervention designs include detailed analysis of and mitigation strategies for political risk 4. Intervention designs include detailed analysis of and mitigation strategies for reputational risk 5. Risks are routinely monitored and reflected upon by the partnership 6. Risk mitigation actions taken by the partnership are documented and communicated	*Common approach (adjusted), Interim Document Reviews, sample of agency risk strategies*
5.5 Intervention designs include the analysis of cross-cutting issues (as defined in KPI2)	1. Intervention design documentation includes the requirement to analyse cross cutting issues 2. Guidelines are available for staff on the implementation of the relevant guidelines 3. Approval procedures require the assessment of the extent to which cross-cutting issues have been integrated in the design 4. Intervention designs include the analysis of gender issues 5. Intervention designs include the analysis of environmental sustainability and climate change issues 6. Intervention designs include the analysis of good governance issues 7. Plans for intervention monitoring and evaluation include attention to cross cutting issues	*Common approach (adjusted), Interim Document Reviews*
5.6 Intervention designs include detailed and realistic measures to ensure sustainability (as defined in KPI 12)	1. Intervention designs include statement of critical aspects of sustainability, including; institutional framework, resources and human capacity, social behaviour, technical developments and trade, as appropriate 2. Key elements of the enabling policy and legal environment that are required to sustain expected benefits from a successful intervention are defined in the design 3. The critical assumptions that underpin sustainability form part of the approved monitoring and evaluation plan 4. Where shifts in policy and legislation will be required these reform processes are	*Interim Document Reviews*

	addressed (within the intervention plan) directly and in a time sensitive manner	
5.7 Institutional procedures (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.) positively support speed of implementation	1. Internal standards are set to track the speed of implementation 2. Organisation benchmarks (internally and externally) its performance on speed of implementation across different operating contexts 3. Evidence that procedural delays have not hindered speed of implementation across interventions reviewed 4. Evidence that any common institutional bottlenecks in speed of implementation identified and actions taken leading to an improvement	*Common approach (adjusted), sample of Interim Document Reviews*
KPI 6: Works in coherent partnerships directed at leveraging and/or ensuring relevance and catalytic use of resources.		
6.1 Planning, programming and approval procedures enable agility in partnerships when conditions change	1. Mechanisms in place to allow programmatic changes and adjustments when conditions change 2. Mechanisms in place to allow the flexible use of programming funds as conditions change (budget revision or similar) 3. Institutional procedures for revisions permit changes to be made at country/regional/HQ level within a limited timeframe (less than three months) 4. Evidence that regular review points between partners support joint identification and interpretation of changes in conditions 5. Evidence that any common institutional bottlenecks in procedures identified and action taken leading to an improvement	*Busan Partnership for effective development cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)* *OECD Development Cooperation Report 2015: Making Partnerships effective coalitions for action.*
6.2 Partnerships based on an explicit statement of comparative advantage e.g. technical knowledge, convening power/partnerships, policy dialogue/advocacy	1. Corporate documentation contains clear and explicit statement on the comparative advantage that the organisation is intending to bring to a given partnership 2. Statement of comparative advantage is linked to clear evidence of organisational capacities and competencies as it relates to the partnership 3. The organisation aligns its resources/competencies to its perceived comparative advantage 4. Evidence that comparative advantage is deployed in partnerships to positive effect	*Interim Document Reviews*
6.3 Clear adherence to	1. Clear statement on set of expectations for how the organisation will seek to deliver on	*Busan Partnership for effective*

the commitment in the Busan Partnership for Effective Development Cooperation on use of country systems	the Busan commitment/QCPR statement (as appropriate) on use of country systems within a given time period 2. Internal processes (in collaboration with partners) to diagnose the condition of country systems 3. Clear procedures for how organisation to respond to address (with partners) concerns identified in country systems 4. Reasons for non-use of country systems clearly and transparently communicated 5. Internal structures and incentives supportive of greater use of country systems 6. Monitoring of the organisation trend on use of country systems and the associated scale of investments being made in strengthening country systems	*development cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)* *Quadrennial Comprehensive Policy Review resolution*
6.4 Strategies or designs identify synergies, to encourage leverage/catalytic use of resources and avoid fragmentation	1. Strategies or designs clearly recognise the importance of synergies and leverage 2. Strategies or designs contain clear statements of how duplication/fragmentation will be avoided based on realistic assessment of comparative advantages 3. Strategies or designs contain clear statement of where an intervention will add the most value to a wider change 4. Strategies or designs contain a clear statement of how leverage will be ensured 5. Strategies or designs contain a clear statement of how resources will be used catalytically to stimulate wider change	*Busan Partnership for effective development cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)* *Interim Document Reviews*
6.5 Key business practices (planning, design, implementation, monitoring and reporting) co-ordinated with other relevant partners (donors, UN agencies, etc.)	1. Evidence that the organisation has participated in joint planning exercises, such as the UNDAF 2. Evidence that the organisation has aligned its programme activities with joint planning instruments, such as UNDAF 3. Evidence that the organisation has participated in opportunities for joint programming where these exist 4. Evidence that the organisation has participated in joint monitoring and reporting processes with key partners (donor, UN, etc.) 5. Evidence of the identification of shared information gaps with partners and strategies developed to address these 6. Evidence of participation in the joint planning, management and delivery of evaluation activities	*Common Approach (adjusted), Interim Document Reviews*
6.6 Key information	1. Information on the organisation's website is easily accessible and current	*Interim Document Reviews*

(analysis, budgeting, management, results etc.) shared with strategic/implementation partners on an ongoing basis	2. The organisation has signed up to the International Aid Transparency Initiative or reports through the OECD-DAC systems 3. Accurate information is available on analysis, budgeting, management and is in line with IATI or OECD-DAC (CRS) guidelines 4. Evidence that partner queries on analysis, budgeting, management and results are responded to in a timely fashion 5. Evidence that information shared is accurate and of good quality	
6.7 Clear standards and procedures for accountability to beneficiaries implemented	1. Explicit statement available on standards and procedures for accountability to beneficiary populations e.g. Accountability to Affected Populations 2. Guidance for staff is available on the implementation of the procedures for accountability to beneficiaries 3. Training has been conducted on the implementation of procedures for accountability to beneficiaries 4. Programming tools explicitly contain the requirement to implement procedures for accountability to beneficiaries 5. Approval mechanisms explicitly include the requirement to assess the extent to which procedures for accountability to beneficiaries will be addressed within the intervention 6. Monitoring and evaluation procedures explicitly include the requirement to assess the extent to which procedures for accountability to beneficiaries have been addressed within the intervention	*IASC Taskforce: Accountability to Affected Populations*
6.8 Participation with national and other partners in mutual assessments of progress in implementing agreed commitments	1. Evidence of participation in joint performance reviews of interventions e.g. joint assessments 2. Evidence of participation in multi-stakeholder dialogue around joint sectoral or normative commitments 3. Evidence of engagement in the production of joint progress statements in the implementation of commitments e.g. joint assessment reports 4. Documentation arising from mutual progress assessments contains clear statement of the organisation’s contribution, agreed by all partners 5. Surveys or other methods applied to assess partner perception of progress	*Busan Partnership for effective development cooperation (Paris Declaration on Aid Effectiveness & Accra Agenda for Action)*

6.9 Deployment of knowledge base to support programming adjustments, policy dialogue and/or advocacy	1. Statement in corporate documentation explicitly recognises the organisation's role in knowledge production 2. Evidence of knowledge products produced and utilised by partners to inform action 3. Knowledge products generated and applied to inform advocacy at country, regional or global level 4. Evidence that knowledge products generated are timely/perceived as timely by partners 5. Evidence that knowledge products are perceived as high quality by partners 6. Evidence that knowledge products are produced in a format that supports their utility to partners	*Interim Document Reviews*
KPI 7: Strong and transparent results focus, explicitly geared to function		
7.1 Leadership ensures application of an organisation-wide RBM approach	1. Corporate commitment to a result culture is made clear in strategic planning documents 2. Clear requirements/incentives in place for the use of an RBM approach in planning and programming 3. Guidance for setting results targets and develop indicators is clear and accessible to all staff 4. Tools and methods for measuring and managing results are available 5. Adequate resources are allocated to the RBM system 6. All relevant staff are trained in RBM approaches and methods	*Common approach* *UNDG Results based management Handbook, 2011* *OIOS Inspection and Evaluation manual. 2014* *OECD workshop on the results agenda 2013*
7.2 Corporate strategies, including country strategies, based on a sound RBM focus and logic	1. Organisation-wide plans and strategies include results frameworks 2. Clear linkages exist between the different layers of the results framework, from project through to country and corporate level 3. An annual report on performance is discussed with the governing bodies 4. Corporate strategies are updated regularly 5. The annual corporate reports show progress over time and notes areas of strong performance as well as deviations between planned and actual results	*Common approach (adjusted)*
7.3 Results targets set based on a sound evidence base and logic	1. Targets and indicators are adequate to capture causal pathways between interventions and the outcomes that contribute to higher order objectives 2. Indicators are relevant to the expected result to enable measurement of the degree of	*Common approach (adjusted)*

	goal achievement 3. Development of baselines are mandatory for new Interventions 4. Results targets are regularly reviewed and adjusted when needed	
7.4 Monitoring systems generate high quality and useful performance data	1. The corporate monitoring system is adequately resourced 2. Monitoring systems generate data at output and outcome level of the results chain 3. Reporting structures are clear 4. Reporting processes ensure timely data for key corporate reporting, and planning 5. A system for ensuring data quality exist 6. Data adequately captures key corporate results 7. Adequate resources are allocated to the monitoring system	*Common approach (adjusted)*
7.5 Performance data transparently applied in planning and decision-making	1. Planning documents are clearly based on performance data 2. Proposed adjustments to interventions are clearly informed by performance data 3. At corporate level, management regularly reviews corporate performance data and makes adjustments as appropriate 4. Performance data support dialogue in partnerships at global, regional and country level	*Common approach (adjusted)* *OECD DAC results workshop*
KPI 8 Evidence-based planning and programming applied		
8.1 A corporate independent evaluation function exists	1. The evaluation function is independent from other management functions such as planning and managing development assistance *(operational independence)* 2. The Head of evaluation reports directly to the Governing Body of the organisation *(Structural independence)* 3. The evaluation office has full discretion in deciding the evaluation programme *4.* A separate budget line (approved by the Governing Body) ensures *budgetary independence* 5. The central evaluation programme is fully funded by core funds 6. Evaluations are submitted directly for consideration at the appropriate level of decision-making pertaining to the subject of evaluation	*Common approach* *UNEG Norms and Standards, OECD DAC Evaluating Development Cooperation, Summary of Key norms and standards, Second Edition* *OIOS Inspection and Evaluation manual*

	7. Evaluators are able to conduct their work throughout the evaluation without undue interference by those involved in implementing the unit of analysis being evaluated. *(Behavioural independence)*	
8.2 Consistent, independent evaluation of results (coverage)	1. An evaluation policy describes the principles to ensure coverage, quality and use of findings, including in decentralised evaluations 2. The policy/an evaluation manual guides the implementation of the different categories of evaluations, such as strategic, thematic, corporate level evaluations, as well as decentralized evaluations 3. A prioritized and funded evaluation plan covering the organisation's planning and budgeting cycle is available 4. The annual evaluation plan presents a systematic and periodic coverage of the organisations' Interventions, reflecting key priorities 5. Evidence from sample countries demonstrate that the policy is being implemented	*UNEG Norms and Standards, OECD DAC Evaluating Development Cooperation, Summary of Key norms and standards, Second Edition OIOS Inspection and Evaluation manual Common approach*
8.3 Systems are applied to ensure the quality of evaluations	1. Evaluations are based on design, planning and implementation processes that are inherently quality oriented 2. Evaluations use appropriate methodologies for data-collection, analysis and interpretation 3. Evaluation reports present in a complete and balanced way the evidence, findings, conclusions, and where relevant, recommendations 4. The methodology presented incudes the methodological limitations and concerns 5. A process exists to ensure the quality of all evaluations, including decentralized evaluations	*UNEG Norms and Standards* *Common approach*
8.4 Mandatory demonstration of the evidence base to design new interventions	1. A formal requirement exists to demonstrate how lessons from past interventions have been taken into account in the design of new interventions 2. Clear feedback loops exist to feed lessons into new interventions design 3. There is evidence that lessons from past interventions have informed new interventions 4. Incentives exist to apply lessons learnt to new interventions 5. The number/share of new operations designs that draw on lessons from evaluative approaches is made public	*UNEG Norms and Standards* *World Bank Corporate Scorecard*
8.5 Poorly performing interventions proactively	1. A system exists to identify poorly performing interventions	*WB Corporate scorecard*

identified, tracked and addressed	2. Regular reporting tracks the status and evolution of poorly performing interventions 3. A process for addressing the poor performance exists, with evidence of its use 4. The process clearly delineates the responsibility to take action	*Common approach*
8.6 Clear accountability system ensures responses and follow-up to and use of evaluation recommendations	1. Evaluation reports include a management response (or has one attached or associated with it) 2. Management responses include an action plan and /or agreement clearly stating responsibilities and accountabilities 3. A timeline for implementation of key recommendations is proposed 4. A system exists to regularly track status of implementation 5. An annual report on the status of use and implementation of evaluation recommendations is made public	*UNEG Norms and Standards and Good Practice Guideline to follow up to evaluations* *Common approach*
8.7 Uptake of lessons learned and best practices from evaluations and other reports	1. A complete and current repository of evaluations and their recommendations is available for use 2. A mechanism for distilling and disseminating lessons learned internally exists 3. A dissemination mechanism to partners, peers and other stakeholders is available and employed 4. A system is available and used to track the uptake of lessons learned 5. Evidence is available that lessons learned and good practices are being applied 6. A corporate policy for Disclosure of information exists and is also applied to evaluations	*UNEG Norms and Standards* *UNEG Good Practice Guidelines to follow up to evaluations* *Common approach*
KPI 9: Achievement of development and humanitarian objectives and results e.g. at the institutional/corporate wide level, at the regional/corporate-wide level and regional/country level, with results contributing to normative and cross-cutting goals.		
NOTE: The rating system for KPIs 9-12 is based on that of the OECD DAC's Development Effectiveness Review, where a rating of Highly Satisfactory, Satisfactory, Unsatisfactory, Highly Unsatisfactory or Not Addressed is provided, based on one of the following classifications. This approach aligns with the scoring and rating system proposed for KPIs 1-8 above.		
9.1 Interventions assessed as having achieved their stated	***4. Highly satisfactory:*** Organisations achieve all or almost all intended significant development, normative and/or humanitarian objectives at the output and outcome level	

development and/or humanitarian objectives and attain expected results	***3. Satisfactory:*** Organisations either achieve at least a majority of stated output and outcome objectives (more than 50% if stated) or the most important of stated output and outcome objectives are achieved ***2. Unsatisfactory:*** Half or less than half of stated output and outcome level objectives are achieved ***1. Highly unsatisfactory:*** Less than half of stated output and outcome objectives have been achieved including one or more very important output and/or outcome level objectives ***0. Not addressed***
9.2 Interventions assessed as having realised the expected positive benefits for target group members	***4. Highly satisfactory:*** Interventions have resulted in widespread and significant positive changes experienced by target group members as measured using either quantitative or qualitative methods. (These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency) ***3. Satisfactory:*** Interventions have resulted in positive changes experienced by target group members (at the individual, household or community level). These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency ***2. Unsatisfactory:*** Interventions have resulted in no or very few positive changes experienced by target group members. These benefits may include the avoidance or reduction of negative effects of a sudden onset or protracted emergency ***1. Highly unsatisfactory:*** Problems in the design or delivery of interventions mean that expected positive benefits for target group members have not occurred or are unlikely to occur ***0. Not addressed***
9.3 Interventions assessed as having contributed to significant changes in national development policies and programmes (policy and capacity impacts), or needed system reforms	***4. Highly satisfactory:*** Interventions have made a substantial contribution to either re-orienting or sustaining effective national policies and programmes in a given sector or area of development disaster preparedness, emergency response or rehabilitation. The supported policies or programmes are expected to result in improved lives of target group members ***3. Satisfactory:*** Interventions have made a substantial contribution to either re-orienting or sustaining effective national policies and programmes in a given sector or area of development disaster preparedness, emergency response or rehabilitation ***2. Unsatisfactory:*** Interventions have not made a significant contribution to the development of national policies and programmes in a given sector or area of development, disaster preparedness, emergency response or rehabilitation. (Policy changes in humanitarian situations may include allowing access to affected populations) ***1. Highly unsatisfactory:*** National policies and programmes in a given sector or area of development (including disaster preparedness, emergency response and rehabilitation) were deficient and required strengthening but interventions have not addressed these ***0. Not addressed***
9.4 Interventions assessed as having helped improve gender equality and the empowerment of	***4. Highly satisfactory:*** Interventions achieve all or nearly all of their stated gender equality objectives ***3 .Satisfactory:*** Interventions achieve a majority (more than 50%) of their stated objectives ***2. Unsatisfactory:*** Interventions either lack gender equality objectives or achieve less than half of their stated gender equality objectives. (Note: where a programme or activity is clearly gender-focused (maternal health programming for example) achievement of more than half

women	its stated objectives warrants a satisfactory ***1. Highly unsatisfactory:*** Interventions are unlikely to contribute to gender equality or may in fact lead to increases in gender inequalities ***0. Not addressed***
9.5 Interventions assessed as having helped improve environmental sustainability/helped tackle the effects of climate change	***4. Highly satisfactory:*** Interventions include substantial planned activities and project design criteria to achieve environmental sustainability and contribute to tackle the effects of climate change. These plans are implemented successfully and the results are environmentally sustainable and contribute to tackling the effects of climate change ***3. Satisfactory:*** Interventions include some planned activities and project design criteria to ensure environmental sustainability and help tackle climate change. These activities are implemented successfully and the results are environmentally sustainable and contribute to tackling the effects of climate change ***2. Unsatisfactory:*** *EITHER:* Interventions do not include planned activities or project design criteria intended to promote environmental sustainability and help tackle the effects of climate change. There is, however, no direct indication that project or program results are not environmentally sustainable. AND/OR: Intervention includes planned activities or project design criteria intended to promote sustainability but these have not been implemented and/or have not been successful ***1. Highly unsatisfactory:*** Interventions do not include planned activities or project design criteria intended to promote environmental sustainability and help tackle climate change. In addition changes resulting from interventions are not environmentally sustainable/do not contribute to tackling climate change. ***0. Not addressed***
9.6 Interventions assessed as having helped improve good governance (as defined in 2.1.c)	***4. Highly satisfactory:*** Interventions include substantial planned activities and project design criteria to promote or ensure 'good governance'. These plans are implemented successfully and the results have helped promote or ensure 'good governance' ***3. Satisfactory:*** Interventions include some planned activities and project design criteria to promote or ensure 'good governance'. These activities are implemented successfully and the results have promoted or ensured 'good governance' ***2. Unsatisfactory:*** *EITHER:* Interventions do not include planned activities or project design criteria intended to promote or ensure 'good governance'. There is, however, no direct indication that project or program results will not promote or ensure 'good governance'. AND/OR: Intervention include planned activities or project design criteria intended to promote or ensure 'good governance' but these have not been implemented and/or have not been successful ***1. Highly unsatisfactory:*** Interventions do not include planned activities or project design criteria intended to promote or ensure 'good governance'. In addition changes resulting from interventions do not promote or ensure 'good governance' ***0. Not addressed***

9.7 Interventions assessed as having helped improve human rights	***4. Highly satisfactory:*** Interventions include substantial planned activities and project design criteria to promote or ensure human rights. These plans are implemented successfully and the results have helped promote or ensure human rights. ***3. Satisfactory:*** Interventions include some planned activities and project design criteria to promote or ensure human rights. These activities are implemented successfully and the results have promoted or ensured human rights. ***2. Unsatisfactory:*** *EITHER:* Interventions do not include planned activities or project design criteria intended to promote or ensure human rights. There is, however, no direct indication that project or program results will not promote or ensure human rights, AND/OR: Intervention include planned activities or project design criteria intended to promote or ensure human rights but these have not been implemented and/or have not been successful ***1. Highly unsatisfactory:*** Interventions do not include planned activities or project design criteria intended to promote or ensure human rights. In addition changes resulting from interventions do not promote or ensure human rights. ***0. Not addressed***
KPI 10: Relevance of interventions to the needs and priorities of partner countries and beneficiaries, and extent to which the organisation works towards results in areas within its mandate	
10.1 Interventions assessed as having responded to the needs /priorities of target groups	***4.Highly satisfactory***: Systematic methods are applied in intervention design (including needs assessment for humanitarian relief operations) to identify target group needs and priorities, including consultation with target groups; and intervention design explicitly responds to the identified needs and priorities ***3.Satisfactory:*** Interventions are designed to take into account the needs of the target group as identified through a situation or problem analysis (including needs assessment for relief operations) and the resulting activities are designed to meet the needs of the target group ***2.Unsatisfactory:*** No systematic analysis of target group needs and priorities took place during intervention design or an some evident mismatch exists between the intervention's activities and outputs and the needs and priorities of the target groups ***1. Highly unsatisfactory:*** Substantial elements of the intervention's activities and outputs were unsuited to the needs and priorities of the target group ***0. Not addressed***
10.2 Interventions assessed as having helped contribute to the realisation of national development goals and objectives	***4. Highly satisfactory:*** Interventions are have played a major role in the achievement of specific national development goals or have contributed to meeting humanitarian relief and recovery objectives agreed to with the national government and/or humanitarian community ***3. Satisfactory:*** Interventions have contributed substantially to the achievement of specific national development goals or have contributed to meeting humanitarian relief objectives agreed to with the national government and/or the humanitarian community ***2. Unsatisfactory:*** EITHER Interventions have contributed only partially to the achievement of specific national development goals or to meeting humanitarian relief objectives agreed to with the humanitarian community AND/OR Interventions have been only partially aligned with the achievement of specific national development goals or the meeting of humanitarian relief objectives agreed to with the

	humanitarian community ***1. Highly unsatisfactory:*** Interventions have not contributed to the achievement of specific national development goals or to meeting humanitarian relief objectives agreed to with the humanitarian community AND/OR Interventions have not been aligned with the achievement of specific national development goals or the meeting of humanitarian relief objectives agreed to with the humanitarian community ***0. Not addressed***
10.3 Results assessed as having been delivered as part of a coherent response to an identified problem	***4. Highly satisfactory:*** The organisation consistently achieved a high level of partnership in implementing its interventions ***3. Satisfactory:*** The organisation has improved the effectiveness of its partnership relationship with partners over time and improvements are noted in evaluations ***2. Unsatisfactory:*** The organisation has experienced significant difficulties in developing an effective relationship with partners, and there has been significant divergence between the priorities of the organisation and its partners ***1. Highly unsatisfactory:*** The organisation experiences significant divergence in priorities from those of its (government, NGO or donor) partners and lacks a strategy or plan which will credibly address the divergence and result in strengthened partnership over time ***0. Not addressed***
KPI 11: Results delivered efficiently	
11.1 Interventions assessed as resource/cost efficient	***4. Highly satisfactory:*** Interventions are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time ***3. Satisfactory:*** Results delivered when compared to the cost of activities and inputs are appropriate even when the program design process did not directly consider alternative program delivery methods and their associated costs ***2. Unsatisfactory:*** Interventions do not have credible, reliable information on the costs of activities and inputs and therefore no data is available on cost/resource efficiency ***1. Highly unsatisfactory:*** Credible information is provided which indicates that interventions are not cost/resource efficient ***0. Not addressed***
11.2 Implementation and results assessed as having been achieved on time (given the context, in the	***4. Highly satisfactory:*** All or nearly all objectives of interventions are achieved on time or, in the case of humanitarian programming, a legitimate explanation for delays in the achievement of some outputs/outcomes ***3. Satisfactory:*** More than half of intended objectives of interventions are achieved on time, and this level is appropriate to the context faced during implementation, particularly for humanitarian interventions.

case of humanitarian programming)	***2. Unsatisfactory:*** Less than half of intended objectives are achieved on time but interventions have been adjusted to take account of difficulties encountered and can be expected to improve the pace of achievement in the future. In the case of humanitarian programming, there was a legitimate explanation for delays ***1. Highly unsatisfactory:*** Less than half of stated objectives of interventions are achieved on time, and there is no credible plan or legitimate explanation identified which would suggest significant improvement in achieving objectives on time ***0. Not addressed***
KPI 12: Sustainability of results	
12.1 Benefits assessed as continuing or likely to continue after project or program completion or there are effective measures to link the humanitarian relief operations, to recovery, resilience eventually, to longer-term developmental results	***4. Highly satisfactory:*** Evaluations assess as likely that the intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction and, eventually, development are credible. Further, they are likely to succeed in securing continuing benefits for target group ***3. Satisfactory:*** Evaluations assess as likely that the intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, the strategic and operational measures to link relief to rehabilitation, reconstruction ***2. Unsatisfactory:*** Evaluations assess a low probability that the intervention will result in continued benefits for the target group after completion. For humanitarian relief operations, efforts to link the relief phase to rehabilitation, reconstruction and, eventually, to development are inadequate. (Note, in some circumstances such linkage may not be possible due to the context of the emergency. If this is stated in the evaluation, a rating of satisfactory is appropriate) ***1. Highly unsatisfactory:*** Evaluations find a very low probability that the program/project will result in continued intended benefits for the target group after project completion. For humanitarian relief operations, evaluations find no strategic or operational measures to link relief, to rehabilitation, reconstruction and, eventually, to development ***0. Not addressed***
12.2 Interventions assessed as having built sufficient institutional and/or community capacity for sustainability, or have been absorbed by government	***4. Highly satisfactory:*** Interventions have contributed to significantly strengthen institutional and/or community capacity as required or institutional partners and communities already had the required capacity to sustain results ***3. Satisfactory:*** Interventions may have contributed to strengthening institutional and/or community capacity but with limited success ***2.Unsatisfactory:*** Interventions have failed to contribute to strengthening institutional and/or community capacity or, where appropriate, to strengthen local capacities for delivery of relief operations and/or for managing the transition to recovery/resilience or development ***1. Highly unsatisfactory:*** Interventions failed to address the need to strengthen institutional and/or community capacity as required. In the case of humanitarian operations, intervention design failed to take account of identified needs to strengthen local capacities for delivery of relief operations and/or for managing the transition to recovery/resilience/development ***0. Not addressed***
12.3 Interventions assessed as having	***4. Highly satisfactory:*** Interventions have made a significant contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for

strengthened the enabling environment for development	participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets. Further, these improvements in the enabling environment are leading to improved development, humanitarian and normative results ***3. Satisfactory:*** Interventions have made a notable contribution to changes in the enabling environment for development including one or more of: the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets ***2. Unsatisfactory:*** Interventions have not made a notable contribution to changes in the enabling environment for development ***1. Highly unsatisfactory:*** For development interventions, there were important weaknesses in the enabling environment for development (the overall framework and process for national development planning; systems and processes for public consultation and for participation by civil society in development planning; governance structures and the rule of law; national and local mechanisms for accountability for public expenditures, service delivery and quality; and necessary improvements to supporting structures such as capital and labour markets). Further, the organisation's interventions failed to address the identified weakness successfully, further limiting results ***0. Not addressed***

Annex D: Checklist for multilateral organisations

During the MOPAN assessment, data is gathered through:

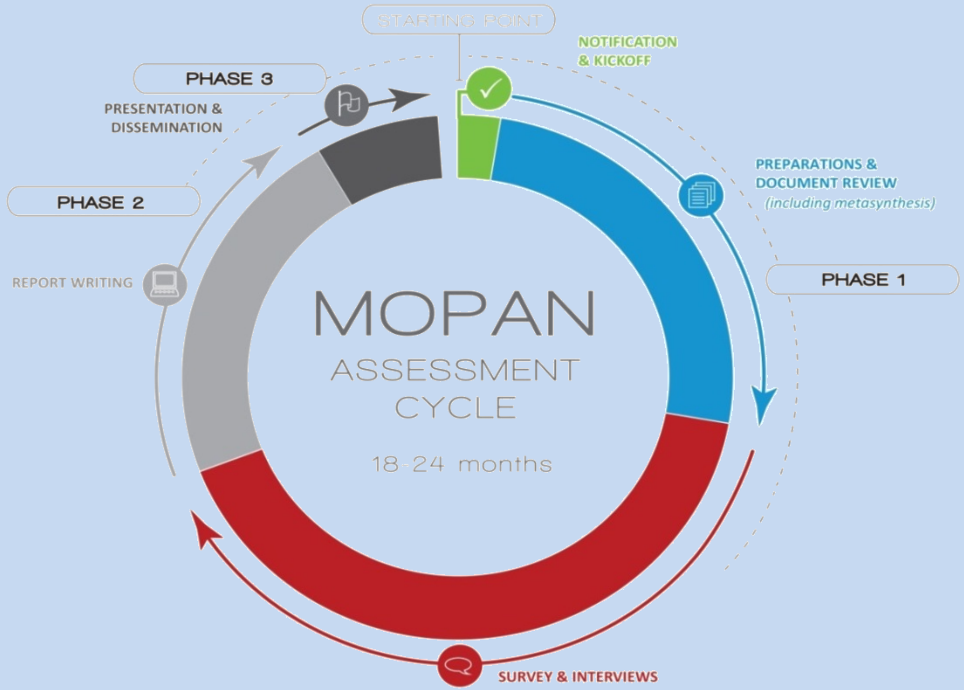
- **document review** of material published and/or commissioned by the multilateral organisation
- **survey** of MOPAN members and other relevant external stakeholders at HQ, capital and country level
- **interviews** with organisation management
- **consultations** with technical staff, MOPAN representatives to the organisation and relevant stakeholders at regional and/or country level

Organisations are assessed against a set of Key Performance Indicators and supporting micro-indicators, which are contextualised to each organisation's operations.

For the purpose of the assessment, the main Network interlocutors are the MOPAN Secretariat and IOD PARC, the external service providers who carry out the independent, external assessments. The process is supported by Institutional Leads – member representatives at HQ and capitol level.

Phase 1: Notification & Kick-off; Preparations & Document Review

- ✓ **Participate in an inception call/meeting** with the MOPAN Secretariat (orientation to the Institutional Leads and IOD PARC will be made at various points)
- ✓ **Support the planning of a 2-day inception visit** for the research team and Secretariat to provide an in-depth overview of the process, discuss potential adaptation of the indicator framework to the assessment and start with data collection (meeting key managerial and technical staff)
- ✓ **Participate in an introductory meeting** with the Institutional Leads and relevant MOPAN members to raise awareness about the assessment process
- ✓ **Review the list of documentation sourced and provide a limited amount of documentation against the Key Performance Indicators** to support the document review process
- ✓ **Fact-check the draft interim document review**
- ✓ **Compile a list of external partner/client respondents to the MOPAN survey at the country level** and provide these to the MOPAN Secretariat



Phase 2: Survey & Interviews; Report Writing

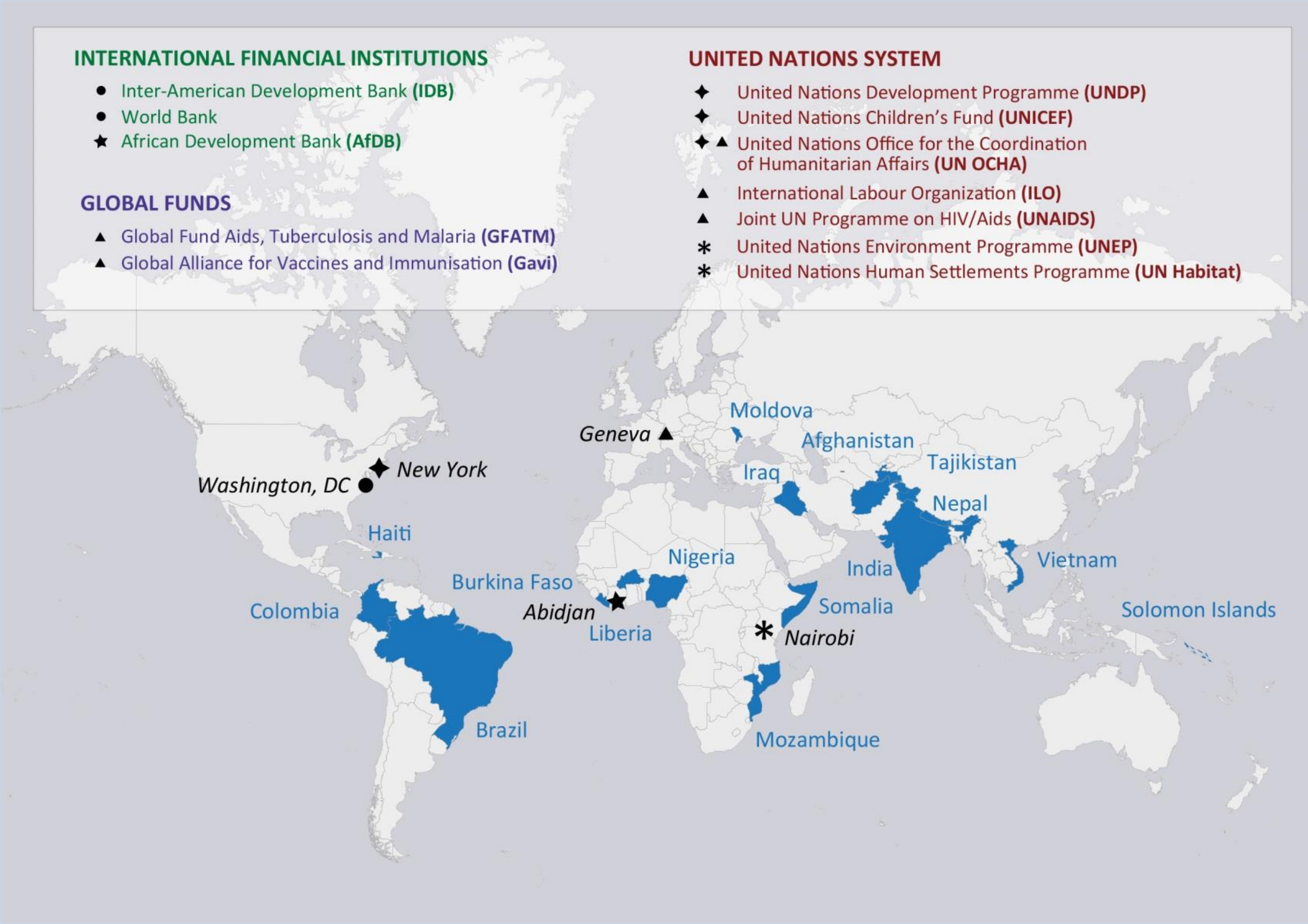
- ✓ Provide information to your survey stakeholders about the MOPAN assessment to build awareness around the exercise, and **send time-series reminders to increase response rates** during the survey process **templates for communication will be provided*

- ✓ **Headquarters interviews** will run over the course of one week. **Schedule and identify relevant managerial and technical staff** for this process; **host the interview team** for one consecutive week (provide one room and internet access for their use) *please do not plan this in the margins of board meetings or other events that will result in absence of key staff from the interview process*
- ✓ **Fact-check** the draft Institutional Report

Phase 3: Finalisation (Preparation & Dissemination)

- ✓ Participate at a senior level in a **presentation of the assessment** by the Institutional Leads and the MOPAN Secretariat
- ✓ Prepare a formal written **Management Response** to the assessment in English and French. This will be published together with the final assessment report.

Annex E: Organisations assessed in 2015-16, and countries where information is collected



Annex F: Organisations assessed in 2017-18, and countries where information is collected