MOPAN 3.0

METHODOLOGY MANUAL
2015-16 ASSESSMENT CYCLE

CONTENTS

ANNEXES

Acronyms

AAPOR	American Association for Public Opinion Research
CO	Country Office
COMPAS	Common Performance Assessment
DER	Development Effectiveness Report
DP	Direct Partner
GFATM	Global Fund to Fight AIDS, Tuberculosis and Malaria
GDI	Gender-related Development Index
GNI	Gross National Income
HDI	Human Development Index
HQ	Headquarters
HR	Human Resources
IDPET	International Programme for Development Evaluation Training
IFI	International Financial Institutions
ILO	International Labor Organisation
IMF	International Monetary Fund
KPI	Key Performance Indicator
MDB	Multilateral Development Bank
MDG	Millennium Development Goals
MI	Micro-Indicator
MOPAN	Multilateral Organisations Performance Assessment Network
OECD-DAC	Organisation for Economic Co-operation and Development Development Assistance Committee
OCHA	Office for the Coordination of Humanitarian Affairs
RBM	Results Based Management
SDG	Sustainable Development Goal
ToR	Terms of Reference
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNEP	United Nations Environment Programme
UNICEF	United Nations Childrens Fund
UN Habitat	United Nations Human Settlements Programme
UNHCR	United Nations High Commissioner for Refugees
UNOCHA	United Nations Office for the Coordination of Humanitarian Affairs

Section 1: Introduction and overview

1. Under MOPAN's reshaped assessment approach, "MOPAN 3.0", the Network is assessing more organisations concurrently than previously (12 over a two-year cycle in 2015-2016, from 4 over a one-year period in 2014), collecting data from more partner countries (16 in 2015-2016), and widening the scope of organisations assessed. MOPAN is also sharpening its focus on results and development effectiveness, while continuing to assess organisational effectiveness.

2. The core hypothesis for 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. The evaluation function of organisations will be examined, and a metasynthesis of evaluation reports will be reviewed to contribute to the development effectiveness component of the MOPAN assessments.

3. MOPAN 3.0 is explicitly not an external audit of an organisation, nor is it an institutional evaluation. As such, MOPAN 3.0 cannot therefore 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 publication akin to that). Nor can MOPAN 3.0 comprehensively document or analyse ongoing organisational reform processes.

4. What MOPAN 3.0 seeks to do is to provide a diagnostic assessment (akin to a snapshot) of information about an organisation at a particular point in time. These assessments seek to identify and examine the journey that an organisation is on, by comparing it only to itself going forward. MOPAN does not compare organisations across the board.

The following series of Operating Principles guides MOPAN 3.0's implementation:

Box 1: Operating Principles

MOPAN 3.0 will generate <u>credible</u>, <u>fair</u> and <u>accurate</u> 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.

---

1. This Methodology Manual describes how these principles will be realised.

2. The MOPAN 3.0 benefit from close involvement of the member governments of the Network, at multiple levels. Figure visualizes the actors involved.

Actors involved in the assessments

3. Each Institutional Lead oversees one organisational assessment. Institutional Leads are MOPAN government representatives, based either in the relevant ministries or at the Permanent Representations/Executive Director offices of the multilateral organisations. At the field level, Country Facilitators – also representatives from the Network members normally based in embassies – support the assessment process in each of the 16 countries. The Network overall is governed by a Steering Committee with representatives from all member countries. The Committee meets two-three times per year to decide on all issues related to the Network, including what organisations to assess and how. The Steering Committee approves all final assessment reports. There is furthermore a Technical Working Group which drives the methodological aspects of the assessments. A permanent Secretariat provides support and co-ordinates the entire process. The Secretariat is since late 2012 hosted by the OECD but reports on all matters of substance to the Steering Committee. An intellectual service provider – IOD PARC – carries out the assessments for the Network, under the guidance of the Technical Working Group and the Secretariat.

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

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

6. *Structure:* This Manual is structured as follows:

    1. Introduction and background
    2. Summary design of MOPAN 3.0
    3. Theory of change
    4. Evidence streams
    5. Indicator framework
    6. Indicative ratings system
    7. Analysis, triangulation and validation
    8. Evidence management
    9. Anticipated strengths and weaknesses
    10. Multilateral organisation and country selection
    11. Reporting
    12. Learning and Engagement

Section 2: Summary Design of MOPAN 3.0

7. **Framing questions***:* Achieving the right approach for MOPAN 3.0 depends on ***asking the right (and timely) questions.*** In the light of the changing context surrounding multilateral effectiveness, the following questions have framed the development of the approach to MOPAN 3.0:

Box 2: Framing questions – MOPAN 3.0

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?**
Are MOs delivering and demonstrating relevant and sustainable results in a cost-efficient way?

---

8. **Theoretical basis:** To address these questions, and in the light of recent theoretical and conceptual work, a Theory of Change has been developed for MOPAN 3.0. The purpose of the Theory of Change is:
    - To provide the conceptual framework to which the empirical approach (methodology) for MOPAN 3.0 assessments is geared
    - To provide a means of testing and validation for the empirical approach (below)

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

Box 3: Hypothesis underlying 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.

---

10. This hypothesis can also be articulated as an 'effectiveness loop', below:

Figure 1: The effectiveness 'loop'
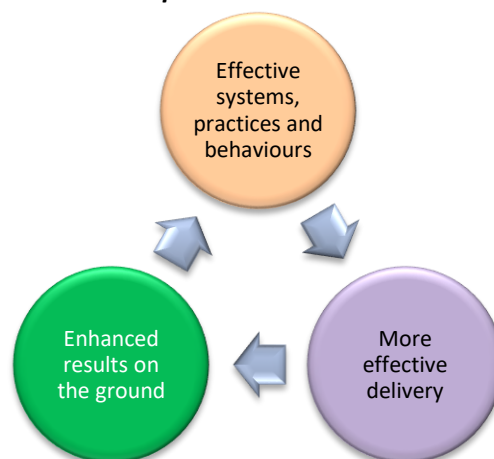
11. As such, 'organisational' and development/humanitarian effectiveness are not perceived under MOPAN 3.0 as discrete performance areas, but rather as a **continuum.**

12. **Performance Areas:** MOPAN 3.0 will assess multilateral organisation performance, through the Indicator Framework, across five Performance Areas. Four of these – Strategic, Operational, Relationship and Performance Management – concern aspects of organisational effectiveness. The Performance Area on Results, addresses development/humanitarian effectiveness.

Figure 2: Performance Areas

STRATEGIC MANAGEMENT *Clear strategic direction, geared to mandate, intended results and realization of global cross-cutting priorities*	RELATIONSHIP MANAGEMENT *Engaging in inclusive partnerships to support relevance, leverage effective solutions and work towards maximum results*
RESULTS *relevant, inclusive, sustainable, efficient*	
OPERATIONAL MANAGEMENT *Organizing assets and capacities behind strategic direction and intended results, to ensure relevance, adaptability and accountability*	PERFORMANCE MANAGEMENT *Systems geared to managing and accounting for development and humanitarian results and the use of performance information, including evaluation and lesson learning*

13. **Empirical design***:* The empirical design of MOPAN 3.0 is grounded in the theoretical approach, and shaped to the Performance Areas, described above. The individual elements of the methodology, and their relation to the theory of change, are set out in Figure 3 below.

Figure 3: Empirical Design
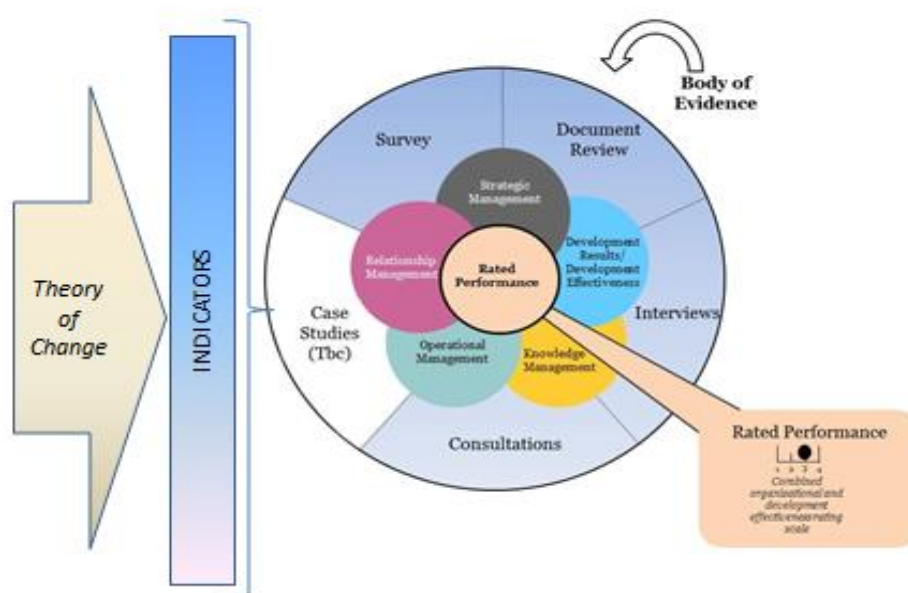
14. The components of this design comprise the following:

   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 through the assessment process.

   ii. ***Lines of evidence.*** Four key lines of evidence (methods) are identified: a document review; a survey; interviews; and consultations. These are extrapolated in Section 4 of this Manual. Possible case studies are also under consideration.

iii. *An indicator framework*, described in Section 5 and presented at Annex 1 (with a comparison of current against former indicators available at Annex 2). This presents the Key Performance Indicators and Micro Indicators against which evidence-based judgements will be made.

iv. *Ratings system*– this will allow for scoring against individual Micro Indicators and Key Performance Indicators. It is discussed in Section 6 of this Manual.

15. **Methodological approach:** The methodology for MOPAN 3.0 will apply four specific approaches. These are:

   i.   A *sequenced* approach
   ii. An *holistic* approach
   iii. A *systematic* approach
   iv. Emphasis on *country/regional level information*

16. These will be operationalised as follows:

17. *A sequenced approach.* This will take place through a 'building blocks' model, with each layer of evidence generated through the sequential assessment process, informed by, and building on, the previous one. Figure 4 explains:

Figure 4: Sequenced approach
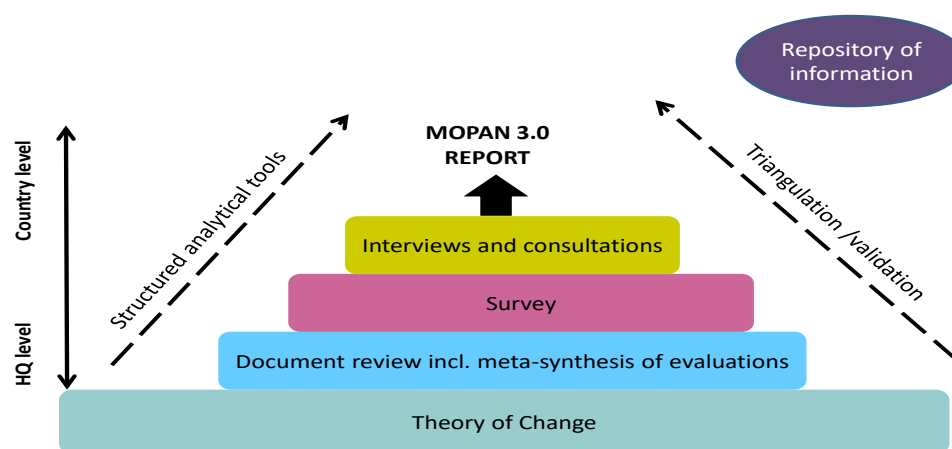
18. Each *layer of information* in the sequence of the assessment process will inform the next, as far as feasibility permits. So the survey, for example, will be informed as far as feasible by findings from the document review, and interviews and consultations will be shaped by findings from both the document review and, if possible, survey.

19. *An holistic approach.* This involves applying *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 at Annex 3 provides a map of intentions here (although these will naturally vary per organisation). This approach will also support triangulation.

20. To assess **development effectiveness:** Evaluations will comprise the primary source of assessment for results (interpreted in multiple forms, as above). Such data is independently generated and itself relies (depending on its quality and credibility) on multiple sources. However, other forms of data will also be included, e.g. management information, to ensure that assessments include sufficiently 'current' information (see for example the description of the document review in Section 4.1).

21. 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) – see following Sections
- For *data analysis:* structured analysis at aggregate level, applying the different lines of evidence within a composite analytical framework, and applying techniques for validation and triangulation. Section 7 of this Manual explains these.

22. Finally, **country/regional level evidence** will be applied against all relevant indicators, through the different data streams including document review and survey information particularly. Analysis/reporting will incorporate this information throughout, in aggregate form, with relevant examples provided where appropriate. Analysis will not result in an overall rating for an organisation's country level performance, but individual data pieces gathered from country level will be collated and stored as part of the assessment (see Section 8 on Evidence Management).

23. The following sections of this Manual describe in detail each of these elements of the methodology; with the hope of presenting, in some methodological detail, a coherent whole.

Section 3: Theory of Change

24. This section of the Methodology Manual describes the Theory of Change for MOPAN 3.0.

25. The Theory of Change is based on the central hypothesis presented in Section 2 above. To reiterate:

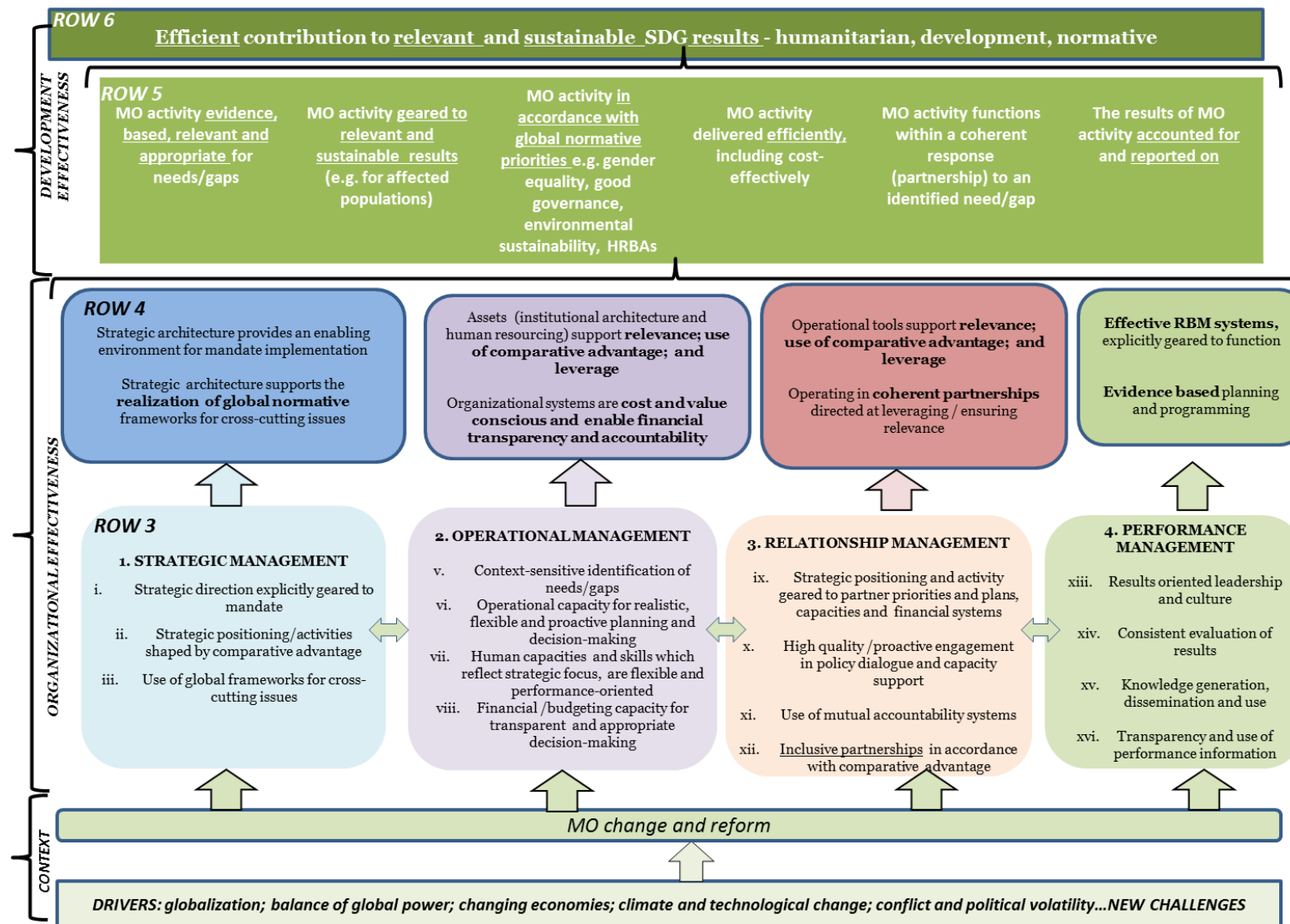
Hypothesis underlying MOPAN 3.0 Theory of Change
- **If** a multilateral organisation has effective systems, practices and behaviours in place *(in terms of strategic, operational, relationship and performance management….)*
  - …**Then** its interventions/activities will be more 'effectively delivered' *(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.

---

26. Guided by a range of conceptual, theoretical and empirical literature, it 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.

27. The Theory of Change is presented below (Figure 5). Whilst requiring specification for each multilateral organisation, particularly in relation to operating structure, it will, in generic form, be interrogated, interpreted and refined through the assessment process. The narrative accompanying it, and the associated assumptions, are described below Figure 5, but 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 5: Theory of Change

ROW 6

Efficient contribution to <u>relevant</u> and <u>sustainable</u> SDG <u>results</u> - humanitarian, development, normative

DEVELOPMENT EFFECTIVENESS

ROW 5

MO activity <u>evidence, based, relevant and appropriate</u> for needs/gaps

MO activity <u>geared to relevant and sustainable results</u> (e.g. for affected populations)

MO activity <u>in accordance with global normative priorities</u> e.g. gender equality, good governance, environmental sustainability, HRBAs

MO activity delivered <u>efficiently,</u> including cost-effectively

MO activity functions within a coherent response (partnership) to an identified need/gap

The results of MO activity <u>accounted for</u> and <u>reported on</u>

ORGANIZATIONAL EFFECTIVENESS

ROW 4

Strategic architecture provides an enabling environment for mandate implementation

Strategic architecture supports the **realization of global normative** frameworks for cross-cutting issues

Assets (institutional architecture and human resourcing) support **relevance; use of comparative advantage; and leverage**

Organizational systems are **cost and value conscious and enable financial transparency and accountability**

Operational tools support **relevance; use of comparative advantage; and leverage**

Operating in **coherent partnerships** directed at leveraging / ensuring relevance

Effective RBM systems, explicitly geared to function

Evidence based planning and programming

ROW 3

1. STRATEGIC MANAGEMENT

i. Strategic direction explicitly geared to mandate
ii. Strategic positioning/activities shaped by comparative advantage
iii. Use of global frameworks for cross-cutting issues

2. OPERATIONAL MANAGEMENT

v. Context-sensitive identification of needs/gaps
vi. Operational capacity for realistic, flexible and proactive planning and decision-making
vii. Human capacities and skills which reflect strategic focus, are flexible and performance-oriented
viii. Financial /budgeting capacity for transparent and appropriate decision-making

3. RELATIONSHIP MANAGEMENT

ix. Strategic positioning and activity geared to partner priorities and plans, capacities and financial systems
x. High quality /proactive engagement in policy dialogue and capacity support
xi. Use of mutual accountability systems
xii. <u>Inclusive partnerships</u> in accordance with comparative advantage

4. PERFORMANCE MANAGEMENT

xiii. Results oriented leadership and culture
xiv. Consistent evaluation of results
xv. Knowledge generation, dissemination and use
xvi. Transparency and use of performance information

CONTEXT

MO change and reform

DRIVERS: globalization; balance of global power; changing economies; climate and technological change; conflict and political volatility...NEW CHALLENGES

9

Narrative of the Theory of Change

28. The following narrative supports the Theory of Change for MOPAN 3.0.

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

30. ***Pathways of connection****:* The logic (pathways of progression) for the MOPAN 3.0 theory of change reflects the approach of organisational and development effectiveness as a continuum.. Specific theorized pathways are as follows:

- **Row 3** posits MOPAN's existing four Performance Areas of organisational effectiveness (with adaptation to that on Performance Management, explained below).
- **Rows 4-6** set out, in varying levels of results hierarchy, key results (whether humanitarian, development or normative) which can, according to the theory, be reasonably 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 (SDGs). It will be assessed through analysis against the Performance Area on Results.

Assumptions

31. 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 the SDGs - will be accomplished.

32. The process of developing the Theory of Change has brought to light a number of key assumptions some of which are also highlighted in the 2013 Independent Evaluation of MOPAN.[3] These will be 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.

---

[2] OECD DAC Development Effectiveness Review
[3] Independent Evaluation of MOPAN. Balogun, P et al (2013)

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

33. All of 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: Evidence Streams

34. 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.

4.1: Document Review

Overview of the Document Review

35. 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.'[1]

36. The Document Reviews are a key plank in the MOPAN methodology, forming the basis on which other evidence streams (survey, interviews and consultations) will rest (see Figure 4 above). A specific Guide to the Document review and Meta-Synthesis of Evaluations has been developed.

Purpose of the Document Review

37. Through an examination of publicly available documents,[2] the Document Review for MOPAN 3.0 will explore two areas:

- Firstly 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

- Secondly, and particularly through a meta-synthesis of evaluations, **evidence of contributions to development, normative or humanitarian results** (development effectiveness)

38. Additionally, and under the sequenced approach above, evidence from the Document Reviews will provide the starting 'evidence block' to inform more focused enquiry under the survey and the interviews/consultations. They will also form a key aspect of triangulation.

Approach

39. The MOPAN methodology includes five Performance Areas, as introduced in Section 2. Under each Performance Area sit a set of Key Performance Indicators (KPIs) (two per Performance Area, other than the Performance Area on Results, which has 4 KPIs). Under each KPI sit a set of Micro Indicators (MIs). It is against these Micro Indicators that evidence will be sought from Document Reviews.

---

[1] Terms of Reference

[2] Documents will be considered to be "publicly available" if they are on the organisation's web site or if the organisation is able to provide them upon request for the purpose of assessing the micro-indicators.

40. Operationally, this implies for analysis under the Document Review the application of a ***single analytical tool,*** geared across the set of Micro Indicators of the methodology, with MIs under KPIs 1-8 reflecting the main dimensions of organisational effectiveness and MIs under KPIs 9-12 focusing on development effectiveness (results). The evidence sources against these groups are anticipated and prioritized (in bold) as follows:

KPIs 1-8	1. **Management information** (organisation's own documentation plus external assessments) 2. Evaluations (independent where feasible)
KPIIs 9-12	3. **Evaluations (independent where feasible)** 4. Management information (though with explicit reporting and requiring a minimum of two different forms of evidence as triangulation)[3] plus external assessments

Step-by-step approach

41. The Document Review, combining management information and meta-synthesis, will adopt the following process. The individual steps and any sub-steps are then set out in the subsequent text.

Figure 6: Structure of the Document Review
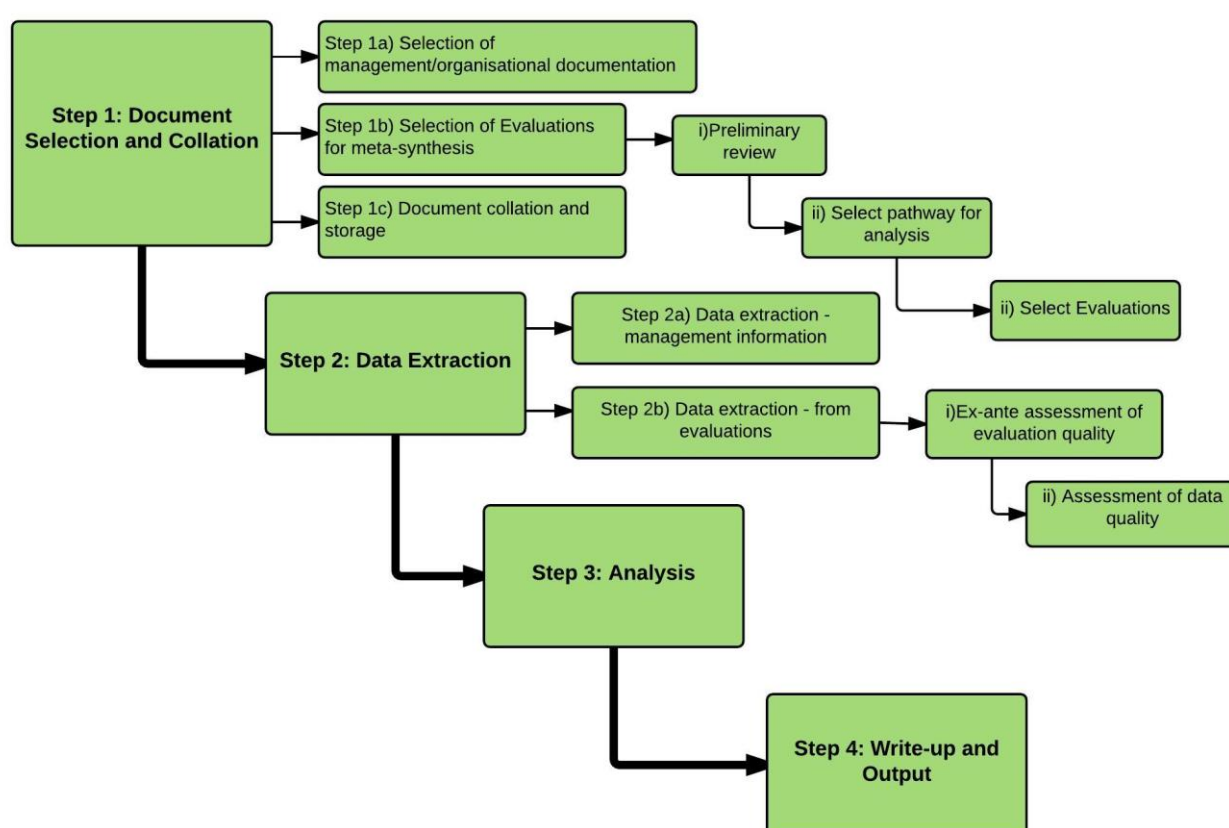
Step 1: Document Selection and Collation

---

[3] In line with the OECD DAC definition of triangulation, available at http://www.oecd.org/dac/2754804.pdf

42. Given the wide range of organisations included for Cycle 1 of MOPAN assessments, there is no 'one size fits all' approach to designing an appropriate selection of documents. However:

- Under the Operating Principle of *efficiency,* (see Section 1) burdens on assessed organisations for document supply need to be reduced
- Under the Operating Principle of *fairness* (see also Section 1), similar types of documentation need to be analysed per organisation
- The volume of assessments under Cycles 1 and 2 require an emphasis on *feasibility.*

43. Document reviews will therefore apply **purposive selection**, aiming to provide coverage of all relevant Micro Indicators, adapted if appropriate to the organisation (see Section 5 on indicator tailoring), guided by a clear and consistent typology. To meet both feasibility and coverage concerns, an intended ceiling of around **35 documents** per organisation will be collated, drawn from a wider screening process and to provide a depth of insight into the organisation's recent and current status. This ceiling will be flexibly applied depending on the scale of the institution. Additionally, up to 15 evaluations will be selected (see step 1b). The following typology will be applied:

Box 4: Indicative typology: Document Review

a: External assessments[4]**,** 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 (see below)

b: 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 (5 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

c: Evaluations
- Independent evaluations

---

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

| | • Organisational evaluations, reviews, or assessments (external or internal) which respond to MIs |

44. The parameters for individual document status are defined in Table 1 below:

Table 1: Parameters for Document Status

Legitimacy	• All documents, regardless of type or level within the organisation, must be approved by the relevant authority[5] (e.g. organisation-wide documents are usually approved by the multilateral organisation's management/Executive Board). • Final versions of evaluations will be included, but not drafts
Accessibility	• Publicly available documentation will be used (information sourced from webpages) • 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 2015-16
	• Strategies, regardless of level within the multilateral organisation, are selected only if they are being implemented in 2015-2016
	• Any information presented on the multilateral organisation's web site is retrieved within 2015-2016 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 the following timelines, unless there is a strong rationale for reviewing older documents:
	   ➤ Project/programme level documents: 2014-2015
	   ➤ Country, regional, or organisation-wide documents: the past three years (since 2012) inclusive of 2013, 2014, 2015 and 2016
	   ➤ Evaluations: from the past year to current (date published: 2014-2015)

Sub-category selection	• Analysis of relevant regional strategies and results frameworks
	• Country level documentation: three types of information:
	      i) country strategy and results framework
	      ii) country strategy evaluation if available (2014-2015)
	      iii) if country strategy evaluation 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.

Step 1 (b) Selection of Evaluations (meta-synthesis)

45. The selection approach for the meta-synthesis of evaluations requires a specific approach. That proposed and applied in the OECD DAC's Development Effectiveness Review (DER) has been adapted, subject to feasibility, for use in MOPAN 3.0, since it provides a thorough and

---

[5] This is intended to follow the same approach as previously, that documents reviewed are final documents (rather than drafts) and that they are providing guidance for organisational behaviour

systematic model, which has been tried and tested with several multilateral organisations.[6] The process will be conducted as follows:

Step 1b (i)

46. Undertake ***preliminary review*** of evaluation universe and quality. This will combine two dimensions:

- Firstly, an assessment of the credibility, including the independence, of the evaluation function;
- Secondly, an assessment of the coverage and quality[7] of development effectiveness reporting.

47. The types of documents to be screened at Preliminary Review stage are set out in Box 5 below:

Box 5: Documents for Preliminary Review – Meta-Synthesis Component
- **Annual Report on Development Results** at the organisational level. For Multilateral Development Banks (MDBs), 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 MDG results across partner countries and may be supplemented (as in the case of UNDP) 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 MDBs 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.
- **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 a UN organisation or to a three to five year strategic plan of a MDB. It may be in a 'Scorecard or Results Framework report' format
- **OECD DAC Peer Reviews** on the evaluation function
- Presence or absence / approach of any **external quality assurance function** for evaluations
- **Selection of:**
  - o Thematic evaluations
  - o Strategic evaluations
  - o Country evaluations
  - o Programme evaluations

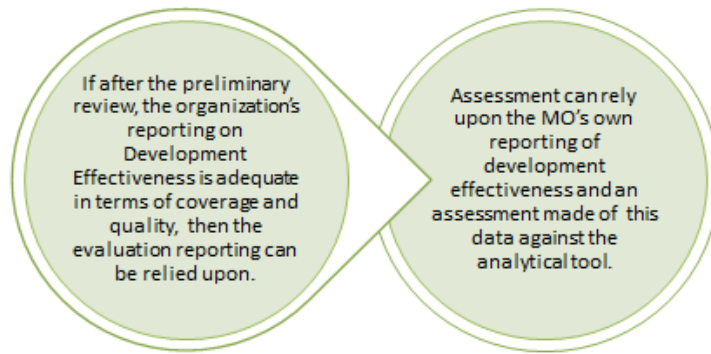*Step 1b (ii) Identify Pathway for analysis*

48. Based on the results of the Preliminary Assessment, a Pathway for Analysis will be enacted. This decision will be made based on the findings of the Preliminary Review, and in conjunction with the relevant Core Team member. Three possible options are available:[8]
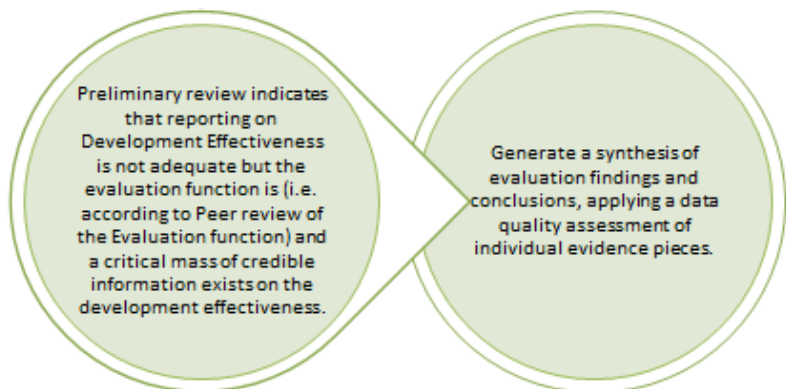
---

[6] Albeit with the caveat that the DER has resourcing of a longer time period and considerably increased resources

[7] As assessed by e.g. an OECD DAC Peer Review, or, if not available, applying the parameters of the Peer Review i.e. coverage, use of evaluation criteria, etc.
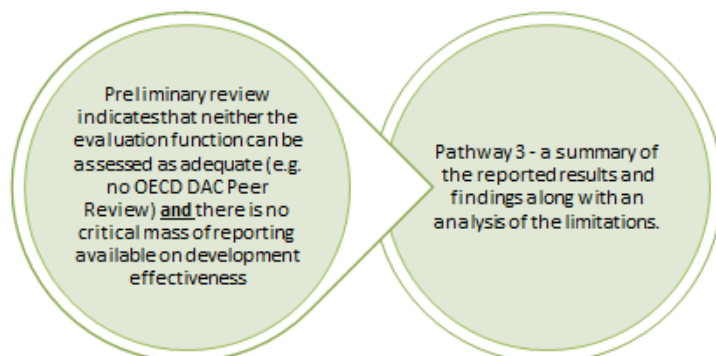
[8] Adapted from DER

If after the preliminary review, the organization's reporting on Development Effectiveness is adequate in terms of coverage and quality, then the evaluation reporting can be relied upon.

Assessment can rely upon the MO's own reporting of development effectiveness and an assessment made of this data against the analytical tool.

Pathway 2: The organisation's reporting on results is not adequate but the evaluation function is:

Preliminary review indicates that reporting on Development Effectiveness is not adequate but the evaluation function is (i.e. according to Peer review of the Evaluation function) and a critical mass of credible information exists on the development effectiveness.

Generate a synthesis of evaluation findings and conclusions, applying a data quality assessment of individual evidence pieces.

Pathway 3: The organisation's humanitarian and development results reporting and available evaluations are inadequate for reporting on development effectiveness

Preliminary review indicates that neither the evaluation function can be assessed as adequate (e.g. no OECD DAC Peer Review) **and** there is no critical mass of reporting available on development effectiveness

Pathway 3 - a summary of the reported results and findings along with an analysis of the limitations.

49. This process will ensure that only credible evidence is included in the meta-synthesis.

Step 1b (iii) Identify the selection

50. Once the appropriate pathway has been selected, relevant reports will be *selected* as follows. The goal for the selection process of Pathways 2 and 3 is to ensure that a reasonable breadth of coverage of agency activities and investments over a specified time-frame are included within the selection.

51.

Table 2: Selection of Results Information

Pathway	Sample documents	
1	As available: • Annual report on development results; • Annual Synthesis of evaluations (focusing on substantive results and reporting) • Annual progress towards Strategic Plan objectives • COMPAS data	Supplemented with: management results reporting under the organisation's strategic plan or similar – but must be triangulated by two other sources e.g. sector reporting, regional/country level reporting etc.
2	Any of the data sources above, if available Sample of: • Policy evaluations • Strategic evaluations • Country evaluations • Thematic evaluations • Project/programme evaluations	
3	Sample of: • Policy evaluations • Strategic evaluations • Country evaluations • Thematic evaluations • Project/programme evaluations	

52. **Where evaluation syntheses are available,** under Pathways 1 and 2, this will mitigate the need to apply individual evaluations. Where individual evaluations require synthesis or summarising, within the resourcing envelope, **up to 15 evaluations reports** will be selected per organisation.

Criteria for selection

53. **Critically, the selection of evaluations does not aim to be representative in terms of coverage** (e.g. of an organisation's financial expenditure per year). Given time pressures, this is unfeasible. Rather a purposive selection will take place, based on agreed parameters.

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

Table 3: Sub-steps for selection of evaluations

Sub-step	Action
1	• Identify the 'evaluation universe', that is to say the number and range of potential evaluation reports for a given Multilateral Organisation that are from the defined time period for inclusion (2014-2015).
2	• Examine a subset of the available evaluation reports to determine how their scope of coverage is defined (if it is) (E.g. in terms of geographic coverage, thematic coverage, objectives coverage, policy areas, technical focus).
3	• Agree with the relevant Core Team Member 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: • 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.
4	• Purposively select up to 15 individual evaluation reports using the criteria defined above.
5	• Agree the scope of coverage with the multilateral organisation, in conjunction with the Institutional Lead and the Secretariat.

Step 1c: Document collation and storage

55. For all elements of the Document Review, including the Meta-Synthesis, **the protocols below will apply to document collation and storage:**

• Identification of the specific documents to be analysed according to the typology above and subject to the indicative ceiling of 35. This will take place through an initial web search for relevant documents available publically 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, mapped per MI, will be prepared (in the form of a bibliography).

• Consultation will then take place with the organisation, via the Institutional Lead/Secretariat, to ensure that the list is complete, and that any additional documents (such as Executive Board Minutes) can be supplied by the organisation

• The list and library will then be finalised

---

[9] Adapted from the DER Sampling process

- Any further requirements as analysis unfolds will be consolidated into a maximum of two further requests, to reduce burdens on the organisation
- If the documents obtained from the third request do not contain the information needed, the consultant team will make the assessment based on the information available
- Documents gathered for the assessment will be stored in a cloud storage system that can be accessed across the assessment team and will be filed systematically according to organisation and relevant document type.

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

Step 2: Data Extraction

57. The first stage in the process will be data extraction – i.e. 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.

58. 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). Analysis will take two forms, depending on the information source, as follows:

Step 2a: Data extraction of management information:

59. 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.

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

- **Quality** – of system, policy, programme or asset e.g. human resources

- **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

- **'Evidence of implementation' –** to ensure a movement beyond 'form' to whether policies and strategies are being implemented in practice

- **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

- **Innovation –** any evidence of innovation (of different types)[10]

61. 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

---

[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.

bibliography will be prepared which includes the full document title, date of publication etc (applying Harvard referencing system).

Step 2b: Data extraction from evaluations

62. For the meta-synthesis component, and for those organisations for whom Pathway 2 or 3 is indicated, **an additional sub-step is required relating the screening of data quality**. This is an important dimension, given the need to include in MOPAN assessments credible and reliable evidence of results.

63. Given the timeline for Cycle 1 of MOPAN Assessments, full ex-ante assessment of evaluation quality for each report is unfeasible. Therefore, a more compressed process will be applied, with individual evaluations, once selected, being assessed for inclusion on two dimensions.

- Firstly, **a short ex-ante assessment** focusing on methodology and credibility of evidence on results

- Secondly, **a data quality assessment** for each relevant piece of evidence.

Step 2b (i) Ex-ante assessment[11]

64. No ex-ante data quality assessment is required under Pathway 1.

65. **Where Pathways 2 or 3 are being followed, a short** (0.5 day) ex-ante review of data quality, focusing specifically on methodology and independence, and explicitly recording assessment against this on a quality scale, will be conducted.

66. An adapted approach of the OECD Development Effectiveness Review will be applied, specifically focused on the DER's second layer of analysis, which applies quality criteria that relate most directly to a review of development effectiveness. [12] Evaluations will be screened and scored for inclusion on the following four parameters:

Evaluation criteria are used	5
Multiple lines of evidence are used. The report indicates that more than one line of evidence (surveys, HQvisits, key informant interviews, document analysis) is used to address the main evaluation issues. One point per line of evidence, to maximum of four.	4
The methodology is well designed. E.g. an explicit theory of how objectives and results were to be achieved, specification of the level of results achieved (output, outcome, impact), baseline data (quantitative or qualitative) on conditions prior to program implementation, a comparison of conditions after program delivery to those before, and	5

---

[11] The process to be undertaken is quality assessment, rather than quality assurance, which would not be appropriate for an external team to conduct under MOPAN.

[12] This is on the basis that:

- Whether or not the evaluation purpose and objectives are stated (dimensions A and B on the DER framework) is important in terms of general evaluation quality, but is not a specific indicator of quality evidence against development and humanitarian results.

- As the DER methodology points out, 'If an evaluation does not score well with respect to the design and implementation of the evaluation methodologies, then it is not likely to provide quality information on…effectiveness.'12

- Whilst the DER does not include the use of evaluation criteria in its second layer of screening, these are considered essential for the analysis of humanitarian and development results.

a qualitative or quantitative comparison of conditions among program participants and those who did not take part.	
Evaluation findings and conclusions are relevant and evidence based. The report includes evaluation findings relevant to the assessment criteria specified. Findings are supported by evidence resulting from the chosen methodologies. There is a clear logical link between the evidence and the findings and conclusions are linked to the evaluation findings as reported.	4
Total	**18**

67. A total of 18 points are available. A minimum threshold of 14/18 points (77%) will be applied for evaluations to meet the threshold for inclusion.

Step 2b (ii) Data quality assessment

68. Data extraction for evaluations will proceed along the same lines as for management information, above. However, an additional step is also required, **under all three pathways,** to ensure the quality and credibility of each individual data piece.

69. A tested approach within evaluation synthesis is to assess the quality of individual evidence pieces for inclusion – whether generated from a cross-section of reliable sources, or whether e.g. reliant on interview data alone. Therefore, as the data extraction process is underway, individual data will be rated along a scale for data reliability and accuracy of:

A	High levels of reliability and accuracy (uses either an internationally recognized source or a combination of multiple sources which are both reliable and accurate)
B	Good levels of reliability and accuracy (uses more than two reliable/accurate source, triangulated)
C	Adequate level of reliability and accuracy (uses one reliable/accurate source with one form of triangulation)
D	Inadequate level of reliability and accuracy (uses only one source which is not demonstrably reliable/accurate e.g. perception data, or a combination of sources which lack reliability/accuracy).

70. Only data categorised as A-C above will be plotted into the analytical template. [13]

71. In terms of evidence content:

- Only findings and conclusions from evaluation reports will be included since recommendations rely heavily on evaluator judgement.

- There are three specific areas where experience suggests that evidence may be weak. These are:

  ➢ **Efficiency,** which often suffers from weak data sources and limited/unreliable/variable methodologies

  ➢ **Sustainability,** which often relies on evaluator judgement and can be variably interpreted (and with varying quality of judgements in terms of realism)

---

[13] This approach has been tested across several global evaluation syntheses and has been found to be effective

> **Impact,** which few evaluation reports other than well-designed specialist impact reports can reliably assess. The indicator framework captures impact under the Performance Area on Results, but does not necessarily require its treatment under the methodology, unless evaluations are available which report on this.

72. As for management information, above, **evidence of development effectiveness under the meta-synthesis** will be plotted into the analytical template, with sourcing identified (referencing to the numbered list of documents generated for the organisation).

Step 3: Analysis

73. For both the wider document review and the meta-synthesis, once evidence has been extracted and plotted into the analytical template, analysis will then need to take place.

74. Document Reviews will not be scored, given the ongoing work to develop a wider ratings system for MOPAN 3.0, which adopts an 'end of process' rather than a sequential approach (see Section 6 for detail on the model of ratings system proposed). Nonetheless, it is important that a layer of analysis is applied to the Document Review both for efficiency reasons and to provide a substantive foundation upon which the successive evidence streams (surveys, interviews and consultations) can be based.

75. 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 disjuncts – 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.

- For both the wider documentary review and the meta-synthesis, 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: Write-up and output

76. An Interim Document Review will be produced for each multilateral organisation. The framework and structure for this will be developed during the analysis phase of the Document Review, but it is not anticipated – for readability – that the Draft Review will report per MI and KPI. Instead, the report will focus on key themes under each relevant KPI. A supporting Annex will identify the main findings per MI and KPI, linked to relevant documentary sources, with clear traceability to sources (so that it is evident which findings relate to the meta-synthesis component, and which to management information).

77. The evaluation meta-synthesis (or summary of results and limitations) will not be separately produced, given time limitations, but will form part of the document review overall.

Standardisation/Consistency

78. Standardisation and consistency will be key parts of the assessment process to ensure the robustness and reliability of assessments. They will be applied throughout the assessment process, as follows:

- Once the Guide to Document Review has been prepared, research analysts within the IOD PARC team (a single analyst per Document Review for the expedited set of four reviews) will be trained on using the analytical template and the methods for data extraction and analysis. As part of training, they will each complete part of one assessment individually.

- For standardisation purposes, analysts and designated IOD PARC Core Team Members per Document Review will then discuss the results and, as a group, reach consensus on the relevant data to be extracted, and how it will be analysed. This process will ensure that analysts and associated Core Team Members have a consistent conceptual understanding of the MIs, and that they are familiar with the types of data sources they are likely to encounter.

- For the evaluation meta-synthesis, the designated research analyst will conduct the relevant ex-ante quality assessment. This will then be peer reviewed by the Deputy Team Leader/methodology specialist.

- 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.

- 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;

- Across the Document Review set, a selection of analysis against MIs will be peer reviewed by the Deputy team leader/Methodology specialist.

- 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. These meetings will be overseen by the Team leader and Deputy team leader.

- Finally, the resulting Interim Document Review will be quality assured by the Core Team member designated to the specific organisation.

Overview

79. The survey seeks to gather both perception data and an understanding of practice from a diverse set of stakeholders of the multilateral organisation under review. For example, whether respondents consider the organisation consistently and independently evaluates its results; or whether or not 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.

80. 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 by 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. A professional survey company will provide support to IOD PARC in the design, implementation and analysis of the survey.

Learning from previous experience of the survey

81. In 2013, MOPAN commissioned an external evaluation of the Common Approach.[17] 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:

- **Quality of data**:

  o A significant percentage of 'don't know' responses which suggests 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 internal operations of the organisation). This undermines validity.

  o The occurrence of respondent fatigue as the survey instrument remains quite long and respondent fatigue may affect both the quality of responses and rate of response.

- **Rating scales**:
  o The rating choices provided in the survey may not be used consistently by all respondents, especially across the many 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.

- **Statistical validity of analytical approach of survey data**:

  o How mean scores are calculated and whether this is valid for category data in the survey as well as how the statistical testing on such data is carried out

82. These concerns and challenges are addressed below.

Respondent types

---

[17] Independent Evaluation of MOPAN. Balogun, P et al (2013) *op. cit.*

83. The approach to the survey under the design of MOPAN 3.0 is to *'ask the right questions to the right people'*. With this in mind, the MOPAN 3.0 survey will target the following primary respondent groups[18]:

- **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.
- **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.

84. 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; though the survey team will retain responsibility for maintaining the respondent list. The total number of respondents will likely vary between assessments.

Respondent selection – first filter

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

86. 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]?
    - o   Less than a year
    - o   At least a year but less than two
    - o   Two years or more

- **Frequency of engagement** – for example: How often does the respondent have contact with [the organisation being assessed]?
    - o   Daily;

---

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

- o Weekly;
- o Monthly;
- o A few times per year or less;
- o Never

Respondent Self-Assessment – Second Filter

87. 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 current 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]?
  - o Not at all familiar
  - o Slightly familiar
  - o Moderately Familiar
  - o Very familiar
  - o Extremely familiar

88. The reason for this choice is that the nature of the questions (some of which are technical and highly detailed) 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 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.[19]

Survey design and customisation

89. 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 particular questions, which are geared to the relevant Micro Indicators of the MOPAN 3.0 indicator framework.

90. 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.

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

Blending Quantitative and Qualitative questions

92. 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

---

[19] E.g. AfDB MOPAN Assessment 2014 Management Response

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.

93. 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.

Administering the survey

94. IOD PARC will ensure the efficient deployment and overall management of the survey instrument. The initial invitation to participate will be sent from the MOPAN contact point at country-level and headquarter respondents or online by the IOD PARC survey team for Direct Partners and any other respondent groups, with an embedded web-link within the email. Senior 'sponsorship' and encouragement has been found to increase response rates and IOD PARC will seek to leverage this, wherever possible, to enhance response rates.

Ethical Standards

95. The survey will be conducted in adherence to international ethical norms and standards, such as those of UNEG[20] and the OECD DAC. Specifically: in order 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

96. 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

97. 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'[21]). Based on a recent meta-analysis[22] 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 nonresponse is unequal among the respondents. Such bias is known as non-response bias (see Survey Design, below, for how this will be mitigated).

---

[20] UNEG - Ethical Guidelines for Evaluation (2008); UNEG Code of Conduct for Evaluation in the UN system
[21] "Response Rates – An Overview." American Association for Public Opinion Research (AAPOR). 29 Sept 2008. http://www.aapor.org/responseratesanoverview
[22] 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

98. 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 will be 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[23].

99. 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.
- Direct partners and any other respondent groups will receive reminders online, issued and tracked by the survey team.

100. 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

101. 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** - 2 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

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

103. Nonresponse bias occurs when some respondents included in the sample do not respond; in essence 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

---

[23] 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

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. In order 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 any day of the week to respond according to their own busy 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.

104. 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

105. The approach to analysing the survey data will need to be revised to align with the reformulated ratings scale, addressed in Section 8 below. The initial thinking on essential aspects of the analysis and treatment of response data is outlined below. Specific detail will be developed as other dimensions of the rating scale are finalised.

First level data analysis

106. The 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'[24] responses and missing data. This analysis also covers the qualitative content analysis of free-text responses to open questions.

107. 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.

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

Weighting scheme

109. The weighting scheme used for the survey responses gives equal weight to the sub-set of responses from each respondent-group/country combination irrespective of how many respondents there were in that sub-set. That would mean that even a small number of, for example, host-government respondents from Country A would contribute the same to the final estimate as a much larger number of, for example, HQ donors responding for Country B.

---

[24] 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.

110. The weighting scheme can be best visualised by imagining the whole set of respondents being divided up onto a grid that has the countries listed across the top (in total, *C* countries) and the respondent groups listed down the side (in total, *R* respondent groups).

	Country 1	Country 2	Country 3	…	Country C
Group 1					
Group 2					
Group 3					
…					
Group R					

111. The number of cells in this table is R x C and therefore required proportional weight for each of these cells is $\frac{1}{RC}$.

112. The total number of respondents is defined as *P* and so the effective number of respondents we wish each cell to contribute is $\frac{P}{RC}$.

113. The number of respondents that falls into any given cell, as we consider it, is defined as G and so we divide the cell's total contribution equally among these respondents in that cell, hence each of the G respondents in that cell is given the same weighting [25] of $W = \frac{P}{RCG}$.

114. These weighted figures should be carefully reviewed, analysed and checked for possibilities of misleading figures due to the number of data points behind them or anomalies in the data.

Second level analysis

115. Procedures for second level analysis will also be confirmed with the development of the final ratings system, to be undertaken from January 2016. Second level analysis considers the difference in responses between sub-groups of the overall respondent data. This will certainly 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.

---

[25] With better mathematical notation, within a population of P respondents, for each of the G respondents falling into country c and respondent r (from a total of C countries and R respondent groups) the weight is: $W_r^c = \frac{P}{RCG_r^c}$

4.3 Interviews and Consultations

116.   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.

4.3.1   Interviews

Approach

117. Interviews form a secondary line of evidence. Nonetheless, these will need to be conducted systematically if the data gathered is a) to have maximum validity, and b) make maximum contribution to the assessments.

118. The aims are:

    i.   To deepen and interrogate the evidence from desk study
    ii. To confirm or clarify trends / findings emerging from document review
    iii. To update on the multilateral organisation's on-going reform agenda and any key changes since documentation was analysed
    iv. To provide contextual insight to clarify, refute and/or validate observations emerging from other lines of evidence/data sources (e.g. document review, survey data)
    v.   To generate new evidence in areas where documentary and survey evidence is lacking
    vi. To seek out explanations and factors ('Why and how' questions)

Sampling

119. Interviewees will be sought at two levels:

- Headquarters level (during an HQ mission, currently scheduled for April-July 2016)
- Country/regional level, during the same time period.

120. *Headquarters:* 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 (see below).

Box 6: 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

---

121. *Country/regional level* At country/regional level, the same principles apply as above, though since interviews will take place remotely, by 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.

122. Possible interviewees at country/regional level are as follows:

Box 7: Interviewees at 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)

---

123. 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.

Methodology

124. 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.

125. Interviews will apply standard ethical protocols,[26] with confidentiality and anonymity assured to interviewees, and without names being either attributed to data or in Annexes. Participants

---

[26] E.g. the UNEG Ethical standards

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

126. 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).

127. 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.

4.3.2 Consultations

128. Consultations are presented in the methodology as discrete from interviews, forming an essentially secondary line of evidence, including contextualisation/ triangulation/ validation/deepening. The assessment team welcomes this distinction, which clarifies and makes explicit the different purposes and objectives of these two often blurred lines of evidence.

Approach

129. Within MOPAN 3.0, consultations, whilst still a significant line of enquiry, will be perceived as having greater scope for customisation and tailoring, with accordingly less structured tools to be applied than for interviews. It is expected accordingly that the criteria for selecting respondents for consultation would also be more flexible than those for interviews, being oriented mostly around gaps / shallow areas in data gathered.

Sampling

130. Consultations will take place at headquarter level only, and are perceived as flexible, depending on 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.

131. Consultations, in the form of group interviews, will therefore be 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

132. 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.

133. 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

134. 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 5: Indicator Framework

135. The four streams of evidence, above, will be systematically collected and geared to an Indicator Framework aligned to the five Performance Areas set out in Section 2 above. The Indicator Framework comprises 8 Key Performance Indicators (KPIs 1-8) aligned to the four Performance Areas on Organisational Effectiveness and four Key Performance Indicators (KPIs 9-12) aligned to the Performance Area on Results.

136. The generic indicator framework for MOPAN 3.0 has the following features.

Box 8: Features of the indicator framework for MOPAN 3.0

- ***Geared to the Theory of Change***, reflecting updated concepts and 'continuum' of 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;
- An emphasis on ***'function' over 'form'***, with the presence of a system, behaviour or practice no longer equating to evidence of 'effectiveness' (rather its application in practice);
- A '***streamlined,'*** approach, with 8 KPIs in total and 46 MIs against organisational effectiveness, and 4 KPIs and 14 MIs against development effectiveness;
- Aiming for ***measurability***, even where indicators are qualitative.

---

137. The Indicator Framework is available at Annex 1.

Adaptation / tailoring of indicators

138. The 'thresholds for adaptation' of indicators is a major concern for MOPAN assessments. 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.

139. This has previously been handled on a case by case basis between organisations under assessment, and assessment teams. MOPAN 3.0 proposes a slightly different approach to respond to the Operating Principles of ***credible, fair and accurate*** assessments, presented in Section 1, above. This will be operationalised as follows:

- Firstly, 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.

- Secondly, whilst adaptation and tailoring under MOPAN 3.0 should therefore be required to a lesser degree than previously, to reflect specificity, 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 frameworks.[27]

- The following process will therefore apply:

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.
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, 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.

140.   Two key nuances to this approach are as follows:

i.   For ***cross-cutting issues:*** these are universal development and humanitarian aims, legitimised and mandated by global frameworks such as the SDGs. Many multilateral organisations work on these issues, although they may not feature explicitly in mandates. They 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,[28] 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 form 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.

ii.   Secondly, a maximum of 5 additional organisation-specific indicators can be considered. However, this must be subject to agreement from the MOPAN Institutional Lead and from the concerned organisation.

141. Given that thresholds for adaptation can be complex, these will also form part of the testing process in a set of four accelerated document reviews to take place from October to December 2015. The four organisations selected reflect diverse organisation types; one International Financial Institution (the World Bank), two UN agencies (UNEP and UNOCHA) and one Global Fund (GFATM)), against which the customisation process will be trialled once the document review is complete.

---

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

[28] 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 6: Indicative Ratings System

Background

142. With evidence in place against the Indicator Framework, derived from the four separate evidence streams discussed in Section 4 above of this Manual, analysis then will take place to provide an overall 'rating' for the assessed organisation. This process will take place during the final stage of the assessment, scheduled for August-October 2016.

143. A ratings system is being developed to achieve:

- The absence of duplication or redundancies between scores (with all findings fed into single set of scores)

- Full complementarity (so no tensions or disjuncts likely to arise between scores across the organisational and development effectiveness components)

- A balance between precise/narrow responses gathered through survey and more complex information from other sources.

Intended approach

144. MOPAN 3.0 will therefore adapt the ratings system during the period of the first cycle of assessments. This will comprise a single, more streamlined scoring and rating system, with a four-point scale, as requested by members.

145. It is important that:

- The broad conceptualisation of the ratings scale is present from the outset, in order that the overall design of the methodology is consistent with its intentions, and not 'out of sync'

- Opportunities for learning /testing are seized as early as possible.

146. Two possible models of rating system have been considered. The first would sequentially score each substantive evidence stream of the assessment – namely, the document review and the survey. This model has been adopted previously by MOPAN. The second model is one more common to evaluative approaches, namely, aggregation of all evidence sources in a single analytical 'box' at the end of the enquiry phase, and application of a rating against the composite evidence base.

147. Both potential models have been given due consideration. However, the former model raises a number of technical and methodological concerns which, on consideration, place it at a disadvantage to the second approach. These relate mostly to the reconciling of two potentially different scores from the document review and the survey, with possible (and even probable) areas of difference/contradiction arising, and therefore likely differences in scoring. Whilst these differences are important – potentially revealing valuable distinctions between HQ level and country level activity – their sequential 'scoring' in isolation, rather than in relation to each other, risks producing numbers which are at best artificial, and at worst a distortion of validity.

148. A further challenge is that, with one source being 'hard' data from documents, and the second being perception data, the issue of relative weighting of scores arises. This is not easily addressed whilst retaining rigour, particularly where the information being rated is not wholly quantitative in nature. Finally, this model also presents a potential implementation challenge,

in that, should these scores be made transparently available, much discussion may focus on the 'number given', distracting from the important business of proceeding with the other evidence streams.

149. The second, more 'evaluative' model will be adopted, where evidence from the multiple data streams is considered in aggregate and in relation to each other at the end of the process, and a 'best fit' rating applied, as described below. This will allow evidence pieces to be viewed in a triangulatory way, with contradictions and tensions identified and treated through the analytical process. It remains important, however, that evidence is fully 'traceable', so that, for example, distinctions between findings from documentary review and perception data (from the survey) are held distinct. The Evidence Density Matrix, attached at Annex 3, provides a statement of intent here, but will need to be carried through in practice to reports.

150. Therefore, under the revision of the ratings system, the evidence from document review and perception data from the survey will **not be separately scored.** However the data from separate evidence streams will be fully traceable within reports; that from document review and potentially, in aggregate form, the survey, will also be available on MOPAN's Repository.

Key features

151. Following consideration of different approaches, the ratings system will draw on the previous 'best fit' model applied by MOPAN under its development effectiveness component. This draws on a set of descriptors of different levels of evidence, and an assessment of where the multilateral organisation 'best fits' within the range of these descriptors.

152. The justification for this type of approach lies mainly in its appropriateness for the MOPAN assessment process – a criteria-referenced basis for judgement is suitable when: criteria are multi-dimensional, there is a mix of both qualitative and quantitative data, and it is not possible to calculate a simple sum of the data points. These conditions apply to the design of MOPAN 3.0.

153. The **best fit model** will be operationalised as follows:

Micro-indicators (MIs):
- MIs will require the assignation of specific criteria (which will be based on existing standards where available e.g. OECD-DAC, UNEG, International Humanitarian Principles etc.) which, when taken together, represent international 'best practice'
- Evidence from the relevant evidence lines applicable to that MI will be applied to the criteria
- Assessment of their presence or absence will generate a rating against the individual MI

154. A four point rating scale will be formulated to a) encompass an emphasis on evidence available, b) emphasise multiple sources (as part of triangulation/verification) and c) reflect the application of criteria as a means of rating per indicator.

Key Performance Indicators (KPIs):
A ratings scale will then be developed for KPIs. However, this requires considerable thought. As discussed above, the new indicator framework has made considerable effort to ensure that component MIs do, in fact, provide sufficient substantive breadth and depth to permit a robust assessment of performance against their aligned KPIs. Nonetheless, whether or not (and how) this approach can be applied to the ratings scale will be a focus for effort in the coming months.

Testing

155. Work is taking place on an ongoing basis to develop selected dimensions of the ratings scale such as the descriptors for the four-point scale and the potential criteria against MIs. Aspects of this will then be tested[29] within the four expedited document reviews, which will also serve as an important means of learning and information for the ratings scale. This should inform the development of the final ratings scale, to be developed and applied to analysis during 2016.

Reporting

156. In reporting, it is particularly important to avoid a reductionist approach – namely, that the rating applied to an organisation does not become the entire focus of the report. A strong narrative will be needed around the rating, both to contextualise this, and to situate the rating on the continuum of change of the organisation. Whilst the ratings scale will tell some of the story, it cannot tell the whole; the supporting narrative, firmly grounded in the analysis of evidence, performs this wider role.

---

[29] Rather than piloted, which implies a more formalized process for which time is lacking

Section 7: Triangulation and Validation

157. As Sections 1-6 of this Methodology Manual have set out, MOPAN 3.0 will apply multiple evidence sources per indicator. Section 6 above describes the use of multiple evidence sources which, where available, will then generate a composite rating.

158. 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

159. Triangulation is the process of using multiple data sources, data collection methods, and/or theories to validate research findings.[30] By using more than a single approach to address a research question, the risk of bias is reduced, and the chances of detecting errors or anomalies increased.[31]

160. Denzin (1978) and Patton (2001) identify four types of triangulation – Methods triangulation; triangulation of sources; analyst triangulation and theory/perspective triangulation.[32] Of these, MOPAN 3.0 will apply three approaches, as follows:

- *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). Secondly, by interrogating data where diverging results arise (e.g. the documentary review shows evidence of a particular policy in place and being used, but survey data indicates little knowledge or use of the policy).
- *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.
- Thirdly, *analyst triangulation -* using multiple analysts to review findings for the documentary review (see above) and composite analytical phases. This will allow for a consistent approach to interpretive analysis.

161. The Evidence Density Matrix at Annex 3 sets out the anticipated evidence sources per MI, to be adapted per assessment. The approach adopted is: At least one 'primary' source of evidence is allocated per Micro Indicator. Where this is not a documentary source, at least two other forms of 'primary' evidence must be available i.e. survey/interview/consultation.

Hypothetical example

---

[30] 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

[31] See for example Morra-Imas, Morr, L and Rist, R (2009), he 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, Carlton University and IOB/Ministry of Foreign Affairs, the Netherlands

[32] 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.

162. To aid comprehension of how analysis will take place, the following **hypothetical** example is provided of an 'evidence box' at the end of the assessment process. *Note that this is a simplified and abbreviated version – in reality there will be many more entries in each box, thus a much richer foundation for the final judgement.*

Box 9: Hypothetical Example of evidence

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

---

From document review

Evidence that the strategic plan and results framework is geared to a long-term vision

Corporate strategy 2013
"At the 2013 Spring Meetings, we adopted two ambitious goals: to end extreme poverty by 2030 and to boost shared prosperity for the poorest 40 per cent in developing countries" (pg. 2)
October 2013: approval of new corporate Strategy, which brought together multiple streams of activity under these two goals.

Key Envisaged Outcomes of the Strategy:
1) "Align all activities and resources to the two goals, maximize development impact, and emphasize the [name of organisation's] comparative advantage.
2) Operationalize the goals through the new country engagement model to help country clients identify and tackle the toughest development challenges
3). Be recognized as a 'solutions' organisation, offering world-class knowledge services and customized development solutions grounded in evidence and focused on results
4). Seek transformational engagements and take smart risks
5). Promote scaled-up partnerships that are strategically aligned with the goals, and crowd in public and private resources, expertise, and ideas
6). Work synergetically, committed to achieving these goals.

[Annual report 2014]

Identification of comparative advantage

"In a world where capital is more easily available, we must emphasize our greatest strengths - the marriage of our vast knowledge with innovative financing to deliver programs that have the greatest impact on the poorest. - i.e. a combining the technical know-how the finance needed to put it to practice.' *[Annual Report 2014 p16]*

From survey:

Responses to question:
"To what extent does the organisation's Country Strategy/intended results in country X reflect its comparative advantage? "

Rate on a scale from 1-4 with 1 being to a large extent and 4 not at all:
- Responses from Government: 55% of respondents rated the organisation's strategy/intended results as reflecting well its comparative advantage
- Responses from partners: 63% of respondents rated the organisation's strategy and intended results as reflecting well its comparative advantage

- Responses from staff: 89% of respondents rated the organisation's strategy and intended results as reflecting well its comparative advantage

From interviews

Quotes from Regional Director from X region

- *"*The organisation has played a key role in this fast growing region because we have been able to provide not only capital but also supported necessary policy changes based on our economic and sector work and policy dialogue. It has however not always been easy as some other donors do not fully agree with our position on agricultural subsidies and the government sometimes claims that our proposed interventions are politically infeasible."

Quotes from government officials from country X

- 'The organisation's strategy is aligned with its comparative advantage and they have clear goals in our country. The problem is that the strategy and goals are difficult to realise, as the organisation either ignores the political aspects of the interventions it proposes or decides to fund activities in a particular region which are only slightly connected to the strategy.'
- 'They do have clear goals…but the problem is that they are heavily influenced by the other actors in the country, so they don't always follow through on what they have committed to do'.

163.   The above example indicates some coherence of evidence against the indicator, in terms of the organisation having a strategy which is aligned with a long term vision/analysis of comparative advantage - but also some contradiction, in that Government and partner respondents, supported by interview evidence, voice concerns about the extent of the alignment. Analysis would therefore recognise this contradiction, and report the evidence transparently, including the facts that a) documentary evidence indicates that, corporately, the organisation's current strategic plan is aligned to clear long-term goals and an analysis of comparative advantage, but that that (in Country X), country-level perception data from partner groups, including government partners, perceived a 'theoretical' element to the organisation's strategic goals and intended results, which was not always a) cognisant of the surrounding political context or b) borne out by its interventions.

Validation

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

- The use of external evaluations and assessments of the organisations to help validate or question the findings on the 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

Purpose

165. Given the emphasis discussed in Section 6 above on traceability of evidence, MOPAN 3.0 has developed an Evidence Management Strategy to transparently set out its intended approaches to managing, storing and recording the evidence it generates. The purpose of the Evidence Management Strategy is twofold:

- 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
- To contribute to the **transparency** of assessments, so that the evidence applied in generating findings is clear and explicit, with the possibility to track back from judgements to evidence.

166. This evidence management strategy describes how evidence will be managed once it has been generated, to realise the purpose above.

Overall approach

167. Evidence generated through MOPAN assessments will have three main forms.

- Firstly, the library of documentation provided for analysis.
- Secondly, completed analytical templates for individual evidence streams. These will be in the following form (Table 4):

Table 4: Forms of evidence

Evidence stream	Template /tool	Methods for validation and triangulation:
Documentary review	Excel spreadsheet geared to each indicator (MI) as tailored (if appropriate) to the specific organisation under assessment	Triangulation – mapping of e.g. data from different documents within the document review, or checking survey responses against each other

Interrogation – checking why tensions/apparent contradictions arise within evidence lines e.g. why respondents from Country Offices may have different opinions on particular issues |
Survey	Systematic recording of response frequency against questions geared to individual indicators (MIs), also as tailored for the specific organisation under assessment	
Interviews	Data recorded on a single analytical template, geared to the indicators, and which a) quantifies proportion of respondents with similar responses and b) provides for explanation/expansion/validation.	
Consultations	As for interviews, above	

168. Thirdly, a **composite analytical template.** This will bring together the key findings, aggregated, from the relevant lines of evidence against a specific MI/KPI – as for example in the Worked Example in Section 7. At an appropriate point, it will also indicate the rating assigned. The template will comprise the main 'running composite' evidence base for the assessment, supported by the discrete pieces of information to be held in the Repository

Managing the evidence base

169. Once the composite evidence base is in place, and ratings per MI and KPI assigned, the following process will be undertaken to ensure transparency (whilst protecting confidentiality):

- **Relevant source data** – where agreed for use in MOPAN assessments – will be collated and stored in the intended Repository. The format of this is yet to be agreed, but its design and management rests with the MOPAN Secretariat.

- The completed **composite analytical templates**, pending screening for confidentiality/organisational sensitivity (below), will be available as an analytical output of the draft assessment report. These can then, if required, be discussed first with the assessed institution and secondly with members as part of validating the assessment report.

- In relation to the above, composite analytical templates will need careful screening to ensure that all data is of a sufficiently aggregated level that confidentiality and anonymity are protected. E.g. that survey data is protected/ the identities of individual interviewees are safeguarded/any internal or confidential data from organisations is neither referenced nor cited / quoted.

170. In terms of reporting:

- Draft reports will be made available to members at the appropriate point (specific procedures and storage mechanisms to be agreed)

- Final reports will also be made available as per agreed procedures.

Section 9: Anticipated Strengths and Weaknesses of the Methodology

171. The expected strengths and limitations of the MOPAN 3.0 methodology will be confirmed and added to as assessments take 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, whilst supported by empirical evidence, will be tested and refined through the assessment process.
- 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.

172. 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.[33] 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.
- The meta-synthesis of evaluations will apply to evaluations generated (i.e approved/ published) in the last two years (2014-2015) in order to apply as up to date information as possible. However, given finalization and 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 effectiveness component of the indicator framework (KPIs

---

[33] 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.

9-12), but this will be explicitly flagged and treated as such, including the triangulation strategies above.

- 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.
- The first four document reviews will support the development and testing of the ratings scale criteria for assessment as mentioned below. Since it is not feasible to develop these ex-ante, they will therefore be a source of testing initial ideas, and learning for further refinement, rather than being able to have the criteria definitively supplied.

Ratings scale

- The ratings scale for MOPAN 3.0 will be developed during the course of 2016. This means that the document reviews, at minimum, will take place in advance of the ratings scale being finalized. Whilst as much work as feasible will take place prior to the document reviews to agree the relevant criteria, these are unlikely to be finalized before the four expedited desk reviews must be delivered. The two processes will therefore operate in tandem with each informing the other.
- With MOPAN 3.0 marking a departure from the past, and MOPAN's methodology having in any event changed considerably since 2009, robust comparisons of 2015-2016 assessments and previous assessments are unlikely to be feasible. However, it may be possible to track the 'changing performance story' within an assessed organisation through the different reports.

Section 10: Multilateral Organisation and Country Selection

Multilateral Organisation Selection

173. 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 the replenishment negotiation process.
- A mix of international financial institutions (IFI), UN funds, programmes, specialised agencies, and humanitarian organisations.

174. In 2015-2016 MOPAN will assess the following twelve organisations: [34]

Box 10: Selected Institutions Cycle 1

International Financial Institutions:
- AfDB
- IDB
- World Bank

Global Funds:
- GAVI
- GFATM

United Nations System:
- ILO
- UNAIDS
- UNDP
- UNEP
- UN Habitat
- UNICEF
- UN OCHA

---

175. 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. A final decision was made at the MOPAN Steering Committee meeting in Washington in May 2015. The four assessments to be expedited are: the World Bank, GFATM, UN OCHA and UNEP.

Country selection

176. MOPAN 3.0 will seek findings from up to 16 countries during the assessment round for Cycle 1. 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.

---

[34] One further organisation is under consideration by members

177. For MOPAN 3.0, the existing criteria for country selection (as applied under previous assessment rounds) have been adapted as follows (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[35]	Retained

---

[35] 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.

178. The adapted parameters were then applied in sequence, as follows (Table 6):

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

179. Specifically, the first stage of the process entailed a mapping exercise, in which the 'presence' of each of the confirmed 12 multilateral organisations in the first cycle was assessed 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.

180. 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.

181. The next stage of the sampling process entailed further criteria-based elimination of countries in order to produce a 'long list' of 24 countries deemed most relevant for assessment. This necessitated the collation 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.

182. These indicators were then systematically scrutinised for each country to develop the preliminary long and short lists of 24 and 16 countries respectively, considered most relevant for country assessments. This was achieved through stratifying the countries according to the above indicators, and prioritising indicators in the following order:

   a. GNI: All countries classified as 'high income' were eliminated, approximately equal representation of low, middle and upper middle income countries was aimed for.
   b. 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 was

given to those displaying higher levels of inequality as measured by the GINI coefficient.

c. Fragile situation: The inclusion of several countries classified as existing in 'fragile situations' by the World Bank was ensured.

d. Governance indicators: Efforts were made to ensure as much representation across each category as was possible, in a bid to capture a variety of operating environments relevant to multilateral co-operation.

183. It was considered particularly 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 above 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 crises.

184. Finally, the proposed list of countries was discussed at a MOPAN Steering Committee meeting in October 2015, and amendments based on member preferences were made. The final list of identified countries (shaded in grey) is displayed on the table below. A more detailed table with the relevant indicators is attached in Annex 4.

Box 11: Countries identified			
Nigeria	Mozambique	Somalia	Liberia
India	Haiti	Colombia	Afghanistan
Vietnam	Nepal	Moldova	Iraq
Brazil	Tajikistan	Solomon Islands	Burkina Faso

Section 11: Reporting

185. Individual institutional reports are produced for each multilateral organisation assessed. In 2014, a synthesis report provided the overall findings and conclusions of the assessment and a two-volume technical report presents details on all micro-indicators and country assessments (Volume I – Data Compendium), and the full methodology and details on the different sources of data that informed the overall findings (Volume II – Methodology Compendium).

186. For 2015, and in response to member requests, a more concise format is proposed, possibly moving away from the existing multi-volume model and towards a less 'linear' presentation of supporting data in particular (although still with a headline Synthesis report providing the key findings and narrative of the assessment).

187. The presentation of country-level data in particular will need to be agreed, since MOPAN 3.0 does not foresee country visits. Therefore, country level data will be subsumed into the wider evidence base for reporting on Organisational and development effectiveness.

188. The precise format and presentation of reports will be agreed with Members during the course of 2015/ early 2016 and prior to embarking on analysis in August 2016; but key principles will be:

- **Accessible** (minimising the use of jargon and ensuring accessible presentation of evidence).
- **Concise** (particularly for the headline Synthesis report, a maximum of 30 pages).
- **Clear** (clear and crystallised narrative for the Synthesis report, with supporting evidence available elsewhere)
- **A 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

189. 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."

190. 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. On the one hand, to enable stronger engagement with the organisations, the Secretariat is developing a formal communication strategy, and on the other, the IOD-PARC team is proposing a learning/engagement strategy to enhance usefulness overall. The strategy will be refined and adjusted as the process unfolds and more experience is gained, and responsibilities will be clarified and allocated.

191. 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 indeed critical if the assessment process and product are to be a spur to organisational reflection and change. The current engagement/organisational learning strategy complements the communication strategy developed and implemented by the Secretariat; and more broadly defines and highlights opportunities to maximize organisational learning and clarify the responsibility of the respective stakeholders.

Opportunities for learning in the design of MOPAN 3.0

192. The design of MOPAN 3.0 presents a range of opportunities for learning. This may occur at both at 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.

193. 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 far as feasible by findings from the document review, and interviews and consultations will be shaped by findings from both 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.

194. 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

195. The mutual learning in the Inception phase focuses on developing a common understanding of scope and methodological issues and various characteristics and aspects of MOPAN's mandate, governance structure, and processes.

196. Engagement has been intense between the Secretariat, the Technical Working Group and the full Steering Committee and IOD Parc.

197. These contacts have provided opportunities to review and absorb lessons from prior MOPAN assessments in view of a changing context, and to reflect on past practice and approaches, and explain the proposed changes and the rationale for these, in a process of mutual learning. They have included the following key issues: presentation by the Secretariat of MOPAN's mission and logic model; a new proposed Theory of Change for the assessment; discussion of the list of organisations to be assessed and the criteria for their selection; review of Terms of Reference for the Institutional Leads; review of other existing work such as the Development Effectiveness Reports and how that work might be used and leveraged; and a preliminary discussion of the existing and proposed new Key Performance Indicators and Micro Indicators, and the rationale for changes.

198. The development, testing and piloting of individual dimensions of the methodology, as they are developed, provide further opportunities for mutual learning across the membership, as does the implementation of surveys. Furthermore, the four "accelerated" document reviews for 2015 (document review methodology including evaluation synthesis, ratings scale and thresholds for adaptation) will be a 'test' and the experience will be thoroughly analyzed and discussed and adjustments made as required.

199. Given that thresholds for adaptation can be complex, these will also form part of the testing process, being trialled in the first accelerated round of assessments, including the document reviews to take place from October to December 2015. The four organisations selected for this accelerated process reflect different types of organisations (one International Financial Institution, two UN agencies and one Global Fund), against which the customization process can be trialled. This is particularly important given the changed approach in MOPAN 3.0 where the revised Indicator Framework, MIs and KPIs have been formulated to explicitly adopt a generic model, rather than being specific to particular organisation types, or applying specified external standards. Strong learning loops on this aspect will therefore be key before the first full round of assessments.

200. Key envisaged IOD-PARC engagement with MOPAN structures is described in the Table 7 below:

Table 7: IOD PARC Engagement with MOPAN Structures

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	Key points during the Inception phase and key reflective points on methodology during implementation
MOPAN Institutional Leads	Preparing for, and organizing, each assessment	Key points during the implementation phase
Country Facilitators	Engagement on application of survey and where appropriate interview process in country	During survey and interview phase
MOPAN Steering Committee	Report approval; possible participation in Steering Committee meetings	Limited other than at report approval stage

Embedding learning in the assessment process

201. There are lessons from evaluation that it is worth keeping in mind for the MOPAN Process. Firstly, the World Bank has recently documented findings on how learning is more likely to come from engagement rather than through the reading of documents[36,] and therefore there must be as much focus on the process as on the final product. Secondly 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] [t]hey 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."[37]

202. Thirdly, the academic and management literature today operates with two aspects of learning: the creation of new knowledge (exploration) and the use of existing knowledge from various sources (exploitation)38. Organisations need to succeed in both exploration and exploitation and keep an 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 our own work as the multiple lines of evidence and the preparation of various assessment tools

---

36 Learning and Results in World Bank Operations: How the Bank Learns, July 2014

37 Patton M. Q, Utilization Focused Evaluation, 2008.

38 J. G. March: Exploration and Exploitation in Organisational Learning, Stanford, 1991

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.

203. In terms of engagement with the organisations, they have in the past been given several opportunities to provide inputs into the assessment process, including to help 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.

204. In terms of engagement with the organisations, the main opportunities are shown in table form below. The first substantive point of engagement is an initial inception call with the organisation, involving the Secretariat, the Institutional Lead, the consultants (IOD PARC) and the Focal Point identified by the organisation. This will set the tone for the assessment process which will run right through to the formal presentation of the final report. The final step is carried out by MOPAN members to the appropriate management and governance entities, as identified in the communication strategy. The Organization Focal Points and the Institutional Leads drawn from the MOPAN membership play important roles at different points in the assessment cycle in terms of helping to create a receptive environment for learning.

Table 8: MOPAN Assessment Cycle: Points of engagement with organisations

Stage of the Assessment cycle	Nature of the engagement	Ways of facilitating the learning element
Notification & Kick Off	Inception Conference Call	Sharing the thinking behind the learning opportunity of MOPAN 3.0
Preparation & Document Review	Introductory Briefing Meeting at Organisation HQ	The Briefing meeting will ideally take place aligned to another event of the organisation thereby facilitating wider awareness and interest on the MOPAN assessment.
	Customisation of the assessment framework	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

205. Notwithstanding the learning that is inherent in the process itself, the final published reports also plays a key role, both in terms of formalizing 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.

206. The presentation of the final report, and the dialogue with the appropriate management and governance level in the assessed organisation is the role of MOPAN members, with the IOD-PARC in a supporting, technical role.

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

208. As Section 11, Reporting, has set out, final reports will, compared to past practice, be more accessible, shorter, less technical and more reader-friendly and will have well evidenced findings that high light both areas for improvement and areas of high performance and good practice in each multilateral organisation.

PERFORMANCE AREA: STRATEGIC MANAGEMENT

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

KPI 1: Organisational architecture[39] and financial framework enables 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 Any other cross-cutting issue included in organisational mandates/commitments

[39] *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: Working in coherent partnerships directed at leveraging / 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, at the regional/country level, and contribution 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)

FACTORS[40] – 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;*

---

[40] 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

68

- 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 multilateral 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

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 2: 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[41] and financial framework enables 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	P	S	P	S

---

41 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
2.1a 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, provide access to justice for all and build effective, accountable and inclusive institutions at all levels)[42]				
2.1d 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		
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

---

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

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
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 programming tools support relevance and agility (within partnerships)				
5.1 Interventions aligned with national/regional priorities and intended national/regional results	P		S	S
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

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
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: Working in coherent partnerships directed at leveraging / 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
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

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
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
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				

KPIs and MIs	Doc. review	Survey	Interviews	Consultations
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, at the regional/country level, and contribution to normative and cross-cutting goals, including Sustainable Development 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
Factors – reasons why results were achieved or not[43] • External – context-related reasons o Operating context o Governance context o Financial context o Partner (national/regional, donor, wider multilateral) context • Internal (signal positive or negative) o Policy issues o Programme or project design; o Objectives /targets – appropriate, realistic o Financial resource issues; o Human resource issues; o Implementation challenges; o Oversight/governance of the institution; o Risk management; o Communication and decision-making systems o Use of innovation (specify)	P

---

[43] 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

KPI 10 Relevance of interventions to the needs and priorities of partner countries and beneficiaries, and extent to which the multilateral 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 3: Shortlisted Countries

This Annex presents the relevant indicators and specific characteristics of sample countries proposed.

Country[44]	MO Density	Income Status[45]	HDI[46]	Inequality Status[47]	GDI[48]	Governance Indicators			ODA
						Political Stability/Absence of Violence[49]	Rule of Law[50]	Corruption[51]	Receipt[52]
Afghanistan*	High (9)	Low ($680)	Low (0.468)	Low (27.82)	Low (0.705)	V. Low	V. Low	V. High	V. High ($5267.46)
Brazil	High (8)	Upper Middle ($11760)	High (0.744)	High (52.67)	Medium (0.411)	Med.	Med.	High	V. High ($1150.06)
Burkina Faso	V. High (10)	Low ($710)	Low (0.388)	Medium (39.78)	Low (0.607)	Low	Med.	High	V. High ($1042.51)
Colombia	V. High (10)	Upper Middle ($7780)	High (0.711)	High (53.53)	Medium (0.460)	Low	Med.	High	High ($851.47)
Haiti*	V. High (10)	Low ($830)	Low (0.471)	Extreme (59.21)	Low (0.599)	Med.	Low	V. High	V. High ($1155.83)
India	V. High (10)	Lower Middle ($1610)	Medium (0.586)	Low (33.9)	Low (0.563)	Low	Med.	High	V. High ($2435.36)

---

44 Countries marked with "*" appear on the World Bank's "Harmonised List of Fragile Situations"

45 https://stats.oecd.org/Index.aspx?DataSetCode=CRS1#

46 http://hdr.undp.org/en/content/table-1-human-development-index-and-its-components

47 http://data.worldbank.org/indicator/SI.POV.GINI

48 http://hdr.undp.org/en/content/table-4-gender-inequality-index

49 http://info.worldbank.org/governance/wgi/index.aspx#reports

50 ibid

51 ibid

52 https://stats.oecd.org/Index.aspx?DataSetCode=CRS1#

Country[44]	MO Density	Income Status[45]	HDI[46]	Inequality Status[47]	GDI[48]	Governance Indicators			ODA Receipt[52]
						Political Stability/Absence of Violence[49]	Rule of Law[50]	Corruption[51]	
Iraq	High (8)	Upper Middle ($6410)	Medium (0.642)	Low (29.54)	Low (0.542)	V. Low	Low	V. High	V. High ($1540.76)
Liberia*	V. High (10)	Low ($400)	Low (0.412)	Medium (38.16)	Low (0.655)	Med.	Low	V. High	High ($533.9)
Moldova	High (9)	Lower Middle ($2550)	Medium (0.663)	Low (30.62)	High (0.306)	Med.	Med.	V. High	High ($346.62)
Mozambique	V. High (10)	Low ($630)	Low (0.393)	High (45.66)	Low (0.657)	Med.	Med.	High	V. High ($2316.38)
Nepal*	V. High (10)	Low ($730)	Low (0.540)	Low (32.82)	Medium (0.479)	Low	Low	V. High	V. High ($2150.56)
Nigeria	V. High (12)	Lower Middle ($2950)	Low (0.504)	Medium (39.74)	N/A	V. Low	Low	V. High	V. High ($2521.62)
Solomon Islands	High (8)	Lower Middle ($1830)	Low (0.491)	N/A	N/A	Med.	Low	High	Med. ($288.32)
Somalia	V. High (11)	N/A	N/A	N/A	N/A	V. Low	V. Low	V. High	High ($987.72)
Tajikistan	High (8)	Lower Middle ($1060)	Low (30.77)	Medium (0.607)	Medium (0.383)	Low	Low	V. High	High ($390.52)
Vietnam	High (9)	Lower Middle $1890)	Medium (0.638)	Medium (35.62)	High(0.322)	Med	Med	High	V. High ($4085.29)

Table 1: Features of proposed countries

Country	Justification for Inclusion	Comparable Alternatives
Africa Africa presented the greatest choice of countries from which to select a purposive sample for the following reason: 1. Africa has 54 countries 2. All of them have a high density of multilateral organisations; this indicator was therefore an insufficient criteria for inclusion 3. Despite considerable contextual variety across countries, this is not fully reflected by conventional development and economic indicators, which are unable to fully capture subtle yet important contextual variations. 'Context of interest' was therefore a crucial factor determining the inclusion of countries in the African selection. e		
Burkina Faso	• Very High MO Density (11) • All round low development indicators, despite very high levels of ODA • Francophone West Africa • Recent political instability: context of 'transition' • Facing severe environmental problems related to desertification; advance of the Sahara	Zambia, Malawi, Mozambique
Liberia	• Very high MO density (10) • Low income status ($400 per capita); the 3rd lowest in Africa. However, Liberia represents a comparatively strong governance context, has made several steps towards economic and social recovery following the conclusion of a damaging civil war in 2003. • Economy has been recently been damaged by a large Ebola outbreak, and MOs have the potential to play an important role in recovery.	Sierra Leone
Mozambique	• Very high MO Density (11) • Low income status ($630 per capita) • High levels of inequality • Low development indicators (HDI, GDI), but relatively strong governance indicators • Pilot country for the UN 'Delivering as One' Approach	Malawi, Zambia, Burkina Faso

Country	Justification for Inclusion	Comparable Alternatives
	• Lusophone Southern Africa	
Nigeria	• Very high MO density (12) • Middle Income Status ($2950 per capita); recently declared Africa's largest economy. Important to capture a middle-income economy in the African selection to provide insights on how MOs respond and perform in this context. • Continues to face numerous development challenges, including an economy heavily dependent on petroleum exports, power shortages, infrastructural deficits, widespread inequality, entrenched regional divisions, corruption and a highly effective insurgency in the North East which has recently attracted significant international attention. • Transitional governance context: although historically susceptible to military intervention in politics, Nigeria has recently concluded its first peaceful transition of power since independence. • Anglophone West Africa •	Kenya, South Africa
Somalia	• Very high MO density (11) • Data deficit: no reliable information pertaining to GDI, HDI, GDI • Extremely low governance indicators: government structures and systems are limited and largely dysfunctional; MOs have a potentially crucial role to play in reconstruction and recovery. 'Fragile Situation'; has experienced profound political unrest and violence for over two decades with numerous regional ramifications.	South Sudan, Democratic Republic of Congo (included in 2014 survey, so reserved for subsequent assessment cycles)
Asia		
MO density in Asia was lower and therefore served as a clearer criterion for country selection, although secondary indicators remained crucial in developing the final Long and Short lists.		
India	• Very high MO density (11) • Lower middle income status ($1610 per capita): not aid-dependen • Continues to face significant development challenges across all sectors	Indonesia, Malaysia

Country	Justification for Inclusion	Comparable Alternatives
Nepal	• Very high MO density (10) • Low income status ($730 per capita) – representation of low income country in the Asian selection • Low HDI scoring • Scene of recent humanitarian disaster requiring significant co-ordination by multilateral agencies	N.A.
Vietnam	• High MO density (9) • Lower middle income status ($1890 per capita) • Medium development indicators • Medium governance indicators	
Afghanistan	• High MO density (9) • Low income status ($680 per capita) • Very low development and governance indicators • 'Fragile Situation': representation of a fragile state in the Asian selection	N.A.
Tajikistan	• High MO density (9) • Lower middle income status, medium HDI and GDI • Representation of CIS region • 'Context of Interest': de-facto one party state	Uzbekistan, Turkmenistan, Kazakhstan, Kyrgyz Republic
The Americas MO density in the Americas was generally low, and this therefore served as the main criterion for inclusion.		
Brazil	• High MO density (8) • Upper middle income status ($11760 per capita); continues to face several developmental challenges, including high inequality, environmental degradation etc. • 'Context of interest': BRIC, also a 'donor' though not within the OECD framework, not aid-dependent.	N.A.
Colombia	• Very high MO density (10); as only one of two countries in the Americas with 'very high' MO density, inclusion was automatic.	N.A.

Country	Justification for Inclusion	Comparable Alternatives
Haiti	• Very high MO density (10); as only one of two countries in the Americas with 'very high' MO density, inclusion was automatic.	N.A.
The Middle East MO density and GNI were the primary criterion for selection, given the relatively low MO density and the presence of countries with high GNI (i.e. Bahrain, Qatar, U.A.E) that were automatically excluded.		
Iraq	• High MO density (8) • Upper middle income status: ($6410 per capita) • 'Fragile Situation': ongoing civil war and resultant humanitarian crisis	N.A.
Europe		
Moldova	• High MO Density (9) • Member request for inclusion	Ukraine
Asia Pacific MO density served as the only criteria for inclusion in the **Pacific**		
Solomon Islands	• High MO Density (8) • Member request for inclusion	

MOPAN 3.0

Scoring and Rating System

2015-16 ASSESSMENT CYCLE

Introduction

This document sets out the Scoring and Rating system to be applied in the MOPAN 3.0 series of assessments. It is based on two previous iterations, which have been reviewed by the MOPAN membership, including its Technical Working Group.

The document sets out the overall framework of the Scoring and Rating system (Section 1), followed by its detailed approach and elements to be applied (Section 2). The Scoring and Rating system will be reviewed following its application in the first round of assessments under MOPAN 3.0, for lessons learned and scope for improvement.

Finally, whilst the scoring and rating is a critical dimension of the MOPAN assessment process, it will be 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.

Section 1: 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 most 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:

Reference point: The approach to Scoring and Rating under MOPAN 3.0 draws on the *OCED's 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 to improve the quality of their indicators.

The approach to be adopted is as follows:

a) *Micro indicator (MI) level*

As in the overall model and Options 1 and 2 above, each of the MOPAN assessment Micro Indicators (MIs) contains 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 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

The below presents a (hypothetical) worked example:

Worked example a):

Example: MI 2.1a on Gender		
Element	**Score**	**Rationale**
Gender equality indicators and targets fully integrated into the organisation's corporate objectives and	4	Gender present in corporate strategic plan and all country and regional plans

strategic plan		
Accountability systems (including corporate reporting and evaluation) reflect gender equality indicators and targets	3	Gender indicators present in corporate plan and all country/regional plans but not within evaluation systems or outputs
Dedicated policy statement on gender equality available and showing evidence of use	2	Policy available but survey data finds limited awareness/use of it at country level
Gender screening checklists or similar tools used for all new initiatives	2	Gender screening tools available but survey finds mixed use of these at country level
Human and financial resources (exceeding benchmarks) are available to address gender issues	3	Gender office exists and is staffed above benchmarks but two staff are consultants (not on permanent contracts)
Capacity development of staff on gender is underway or has been conducted	2	Training has covered some but not all country, regional and HQ offices
TOTAL	16	
AVERAGE PER MI	2.66	
OVERALL MI RATING	**SATISFACTORY**	

b) Aggregation to KPI level

The same logic will be pursued at aggregation to KPI level, to ensure a consistent approach. Taking the average of the constituent scores per MI, a rating is then calculated per KPI. The ratings scale applied will be the same as for MIs, namely:

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

Worked example b)

KPI 5: Operational planning and intervention support tools support relevance and agility (within partnerships).		
MI	Scores per element	Score
5.1	2,3,3	2.66
5.2	2,1,1,2,3,2	1.83
5.3	2,1,1,2,1	1.40
5.4	3,2,2,1,4,3	2.50
5.5	4,3,3,2,2,2,2	2.57
5.6	2,2,1,1	1.50

5.7	1,2,2,2	1.75
TOTAL		14.21
AVERAGE PER KPI		2.03
OVERALL KPI RATING		**SATISFACTORY**

Section 2 Elements of 'best practice' for Micro Indicators

The following presents the elements of best practice agreed by MOPAN members as the basis for assessment against the Micro Indicators (MIs) for MOPAN 3.0.

The tables below set out the proposed elements 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
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 Interim Document Reviews has been applied, alongside a sample of relevant items produced by agencies and, as appropriate, professional knowledge.

KPI 1: Organisational architecture[1] and financial framework enables 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,*

[1] 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 by Gary N. Mc Lean
1.3 The 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. The strategic plan is aligned to wider normative frameworks and associated results 2. The strategic plan includes clear results for normative frameworks 3. A system to track results is in place and being applied 4. Clear accountability is established for achievement of normative results 5. Progress on implementation on an aggregated level is published at least annually	*Common approach*
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	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	*UNSWAP; Gender evaluations (e.g. Joint evaluation of*

women.	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	*Joint Gender Programmes in the UN system).* *Common approach*
2.1b Environmental sustainability and climate change	1. Dedicated policy statement on environmental sustainability and climate change available 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	*UNDP/UNEP: Mainstreaming 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	*World Bank Governance assessment framework; UNDP*

	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	*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*

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. Organisational structure is aligned with, or being reorganized to fit the requirements of, the current Strategic Plan 2. Staffing is aligned with, or being reorganized to, requirements set out in the current Strategic Plan 3. Resource allocations across functions are aligned to current organisational priorities and goals, as set out in the current Strategic Plan 4. 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	*Sample resource mobilization strategies (UNHCR, UNICEF, WFP)*

		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	
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	*Common Approach (with adjustment), sample of UN agency Resource Mobilization Strategies, Interim Document Reviews*
	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	
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	*Common Approach (with adjustment), Interim Document Reviews*
	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	

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 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)	*Common Approach (with adjustment), sample of agency Results Based Budgeting statements*
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 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	*Common Approach (with adjustment), Interim Document Reviews*
KPI 5: Operational planning and intervention support tools support relevance and agility (within partnerships).		
5.1 Interventions aligned with national/regional priorities and intended	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	*Busan Partnership for effective development cooperation (Paris*

national/regional results	3.	Structures and incentives in place for technical staff that allow investment of time and effort in alignment process	*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. 2. 3. 4. 5. 6.	Intervention designs contain a clear statement that positions the intervention within the operating context Context statement has been developed jointly with partners Context analysis contains reference to gender issues, where relevant Context analysis contains reference to environmental sustainability and climate change issues, where relevant Context analysis contains reference to governance issues, including conflict and fragility, where relevant 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 employed	1. 2. 3. 4. 5.	Intervention designs contain a clear statement of capacities of key national implementing partners Capacity analysis considers resources, strategy, culture, staff, systems and processes, structure and performance Capacity analysis statement has been developed jointly where feasible Capacity analysis statement includes clear strategies for addressing any weaknesses, with a view to sustainability 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	*Interim Document Reviews*
5.4 Detailed risk (strategic, political,	1.	Intervention designs include detailed analysis of and mitigation strategies for operational risk	*Common approach (adjusted), Interim*

reputational, operational) management strategies ensure the identification, mitigation, monitoring and reporting of risks	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	*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 addressed (within the intervention plan) directly and in a time sensitive manner	*Interim Document Reviews*
5.7 Institutional	1. Internal standards are set to track the speed of implementation	*Common approach*

procedures (including systems for engaging staff, procuring project inputs, disbursing payment, logistical arrangements etc.) positively support 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	*(adjusted), sample of Interim Document Reviews*
KPI 6: Working in coherent partnerships directed at leveraging/ 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. Evidence that resources/ competencies needed for intervention area(s) are aligned to the perceived comparative advantage 4. Comparative advantage is reflected in the resources (people, information, knowledge, physical resources, networks) that each partner is able (and willing) to bring to the	*Interim Document Reviews*

	partnership	
6.3 Clear adherence to the commitment in the Busan Partnership for Effective Development Cooperation on use of country systems	1. Clear statement on set of expectations for how the organisation will seek to deliver on 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	*Busan Partnership for effective 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,	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	*Common Approach (adjusted), Interim Document Reviews*

UN agencies, etc.)	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	
6.6 Key information (analysis, budgeting, management, results etc.) shared with strategic/implementation partners on an ongoing basis	1.	Clear corporate statement on transparency of information	*Interim Document Reviews*
	2.	The organisation has signed up to the International Aid Transparency Initiative	
	3.	Information is available on analysis, budgeting, management in line with the guidance provided by the International Aid Transparency Initiative	
	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	*IASC Taskforce: 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	
6.8 Participation with	1.	Evidence of participation in joint performance reviews of interventions e.g. joint	*Busan Partnership*

national and other partners in mutual assessments of progress in implementing agreed commitments	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	*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 goal achievement 3. Development of baselines are mandatory for new Interventions 4. Results targets are regularly reviewed and adjusted when needed	*Common approach (adjusted)*
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 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)*	*Common approach* *UNEG Norms and Standards, OECD DAC Evaluating Development Cooperation, Summary of Key norms and standards, Second Edition* *OIOS Inspection and Evaluation manual*
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	*UNEG Norms and Standards, OECD DAC Evaluating Development Cooperation, Summary of Key norms and*

	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	*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 identified, tracked and addressed	1. A system exists to identify poorly performing interventions 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	*WB Corporate scorecard* *Common approach*
8.6 Clear accountability	1. Evaluation reports include a management response (or has one attached or	*UNEG Norms and*

system ensures responses and follow-up to and use of evaluation recommendations	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	*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/country level, and contribution 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 development and/or humanitarian objectives and attain expected results	1. ***Highly satisfactory:*** Organisations achieve all or almost all intended significant development, normative and/or humanitarian objectives at the output and outcome level 2. ***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 3. ***Unsatisfactory:*** Half or less than half of stated output and outcome level objectives are achieved 4. ***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

	5. ***Not addressed***
9.2 Interventions assessed as having realised the expected positive benefits for target group members	1. ***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) 2. ***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 3. ***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 4. ***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 5. ***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	1. ***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 2. ***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 3. ***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) 4. ***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 5. ***Not addressed***

9.4 Interventions assessed as having helped improve gender equality and the empowerment of women	1. ***Highly satisfactory:*** Interventions achieve all or nearly all of their stated gender equality objectives 2. ***Satisfactory:*** Interventions achieve a majority (more than 50%) of their stated objectives 3. ***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 its stated objectives warrants a satisfactory 4. ***Highly unsatisfactory:*** Interventions are unlikely to contribute to gender equality or may in fact lead to increases in gender inequalities 5. ***Not addressed***
9.5 Interventions assessed as having helped improve environmental sustainability/helped tackle the effects of climate change	1. ***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 2. ***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 3. ***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 4. ***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. 5. ***Not addressed***
9.6 Interventions assessed as having helped improve good governance (as defined in 2.1.c)	1. ***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' 2. ***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'
	3. ***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
	4. ***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'
	5. ***Not addressed***

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

10.1 Interventions assessed as having responded to the needs / priorities of target groups	1. ***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
	2. ***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
	3. ***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
	4. ***Highly unsatisfactory:*** Substantial elements of the intervention's activities and outputs were unsuited to the needs and priorities of the target group
	5. ***Not addressed***
10.2 Interventions assessed as having helped contribute to the realisation of national development goals and	1. ***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
	2. ***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

objectives	government and/or the humanitarian community
	3. *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
	4. *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
	5. *Not addressed*
10.3 Results assessed as having been delivered as part of a coherent response to an identified problem	1. *Highly satisfactory:* The organisation consistently achieved a high level of partnership in implementing its interventions
	2. *Satisfactory:* The organisation has improved the effectiveness of its partnership relationship with partners over time and improvements are noted in evaluations
	3. *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
	4. *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
	5. *Not addressed*
KPI 11: Results delivered efficiently	
11.1 Interventions assessed as resource/cost efficient	1. *Highly satisfactory:* Interventions are designed to include activities and inputs that produce outputs in the most cost/resource efficient manner available at the time
	2. *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
	3. *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 4. **Highly unsatisfactory:** Credible information is provided which indicates that interventions are not cost/resource efficient 5. **Not addressed**
11.2 Implementation and results assessed as having been achieved on time (given the context, in the case of humanitarian programming)	1. **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 2. **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. 3. **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 4. **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 5. **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	1. **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 2. **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 3. **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)

	4. **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 5. **Not addressed**
12.2 Interventions assessed as having built sufficient institutional and/or community capacity for sustainability, or have been absorbed by government	1. **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 2. **Satisfactory:** Interventions may have contributed to strengthening institutional and/or community capacity but with limited success 3. **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 4. **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 5. **Not addressed**
12.3 Interventions assessed as having strengthened the enabling environment for development	1. **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 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 2. **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

<table>
<tr>
<td></td>
<td>

as capital and labour markets
3. *Unsatisfactory:* Interventions have not made a notable contribution to changes in the enabling environment for development
4. *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
5. *Not addressed*

</td>
</tr>
</table>