

Draft for review and inputs

APPROACHES TO EVALUATING THE EFFECTIVENESS OF MEASURES TAKEN TO IMPLEMENT THE CONVENTION ON BIOLOGICAL DIVERSITY

I. BACKGROUND

1. Article 26 of the Convention on Biological Diversity calls upon Parties to report on measures taken to implement the Convention and their effectiveness. The main mechanism for reporting under the Convention is the national report. In section II of the sixth national report template, adopted in decision XIII/27, the Conference of the Parties called on Parties to report on the measures taken to implement their national biodiversity strategy and action plan as well as to assess the effectiveness of these. In doing so the COP also requested Parties to include information on the methodologies and tools used to arrive at these assessments.

2. Further in decision XIII/1 the Conference of the Parties encouraged Parties to undertake evaluations of the effectiveness of measures undertaken to implement the Strategic Plan for Biodiversity 2011-2020, to document experiences, including the methodologies applied, to identify lessons learned, and to provide this information to the Executive Secretary, including through their sixth national report and the clearing-house mechanism. Information on the evaluation of effectiveness submitted by Parties through their sixth national reports and other channels will be compiled, as per the request to the Executive Secretariat in decision XIII/27, so that it can be considered at a meeting of SBSTTA or SBI after the deadline for the sixth national reports, i.e. in 2019 or 2020.

3. The current note is intended to provide guidance to Parties and support their efforts in undertaking assessments of the effectiveness of measures they have taken to implement the Convention and to report on these in the sixth national report. Section II of this note examines relevant work on effectiveness already undertaken under the Convention on Biological Diversity, including documents previously prepared for meetings of the Subsidiary Body on Scientific, Technical and Technological Advice, and the Ad Hoc Open-ended Working Group on the Review of Implementation of the Convention.¹ Section III, based on academic literature, presents information on different approaches to evaluating the effectiveness of measures. Section IV contains conclusions. Suggested recommendations for consideration by SBSTTA-21 will be added at a later stage.

Please note: Section III of this draft provides brief descriptions of the types of approaches Parties might consider using as they undertake assessments of the effectiveness of measures taken. It is intended that the document available for SBSTTA-21 would include concrete examples provided by Parties in response to the request in decision XIII/1 for each approach described in this document. These examples would help to illustrate the somewhat academic descriptions in Section III and thereby facilitate the work of Parties as they plan and carry out assessment of the effectiveness of measures taken towards national targets established in the updated NBSAPs. In reviewing this draft, **Parties are therefore invited to share experiences and examples of effectiveness assessments that have previously been undertaken. These would be drawn upon in finalizing the document for SBSTTA-21.**

II. RELEVANT WORK UNDER THE CONVENTION

4. Assessing the effectiveness of measures taken is a process of critical evaluation. Its purpose is to reflect upon the suitability of the measure that has been taken, collect information on how it has been implemented and on its overall effects. The information generated from such an assessment allows for the

¹ See documents UNEP/CBD/SBSTTA/17/3 (Assessing the effects of the types of measures taken in accordance with the provisions of the convention), UNEP/CBD/SBSTTA/19/4 (Tools to evaluate the effectiveness of policy instruments for the implementation of the strategic plan for biodiversity 2011-2020) and UNEP/CBD/WGRI/5/INF/20 (Methodologies of self-assessment by parties of the implementation of the convention).

design and implementation of future measures to be improved as well as to demonstrate the overall impact of the measure in the form in which it has been implemented. Such assessments are particularly important considering the limited financial resources available for implementing biodiversity measures. Assessments allow for these limited resources to be put to best use and can help to generate resources by highlighting successful examples of implementation.

5. Information collected through various processes under the Convention on Biological Diversity indicates that many Parties are evaluating the effectiveness of the measures they are taking to implement their national biodiversity strategy of action plan, the Strategic Plan for Biodiversity 2011-2020 or other measures to implement the Convention. However, specific information on how this is being carried out is limited. For example, of 179 fifth national reports approximately 45 per cent contained explicit assessments of progress towards the Aichi Targets. However these assessments have rarely indicated a critical analysis of the effectiveness of the measures taken to bring about the observed changes.² As such the information in the fifth national reports provides insufficient information to be able to determine how widespread the use of effectiveness assessment is or to determine what approaches Parties have found useful or to draw conclusions and share any lessons learned.

6. The situation is similar with regard to the national biodiversity strategies and action plans. Of the Parties that have submitted a post-2010 NBSAP, 70 mention having done an assessment of their previous NBSAP as part of, or contributing to the revision process. However, the specific methods used to assess these is not usually clear nor is the extent to which they look at issues related to effectiveness of the actions that were planned and carried out as part of NBSAP implementation.³

7. In a survey carried out in 2015⁴, 30 Parties indicated that they had assessed the effectiveness of the actions or interventions undertaken to implement their national biodiversity strategies and action plans or related strategies.⁵ However, the specific methods used to evaluate effectiveness were not generally specified. Building on the results of this survey, a document examining the tools used by Parties to evaluate the effectiveness of policy instruments was prepared for the nineteenth meeting of the Subsidiary Body on Scientific, Technical and Technological Advice. This document outlines in general terms, a variety of approaches being used by countries.⁶

8. A notable exception to the general absence of information on approaches is the widespread use of methods to assess effectiveness with regard to the management of protected areas. A variety of tools to assess the management effectiveness of protected areas are used by Parties including: the protected area management effectiveness (PAME) assessment methodologies; the Rapid Assessment and Prioritization of Protected Area Management (RAPAM) tool; and the Management Effectiveness Tracking Tool (METT), developed by the World Bank and World Wildlife Fund for Nature and the Integrated Management Effectiveness Tool developed in the frame of the EU funded Biodiversity and Protected Areas Management (BIOPAMA) programme. For example as of January 2015, the Global Database on Protected Areas Management Effectiveness had collected 17,739 protected areas management effectiveness assessments, representing 9,037 protected areas. As of a 2015 assessment of submissions to the database, some 42 CBD Parties (or 21.4%) have implemented management effectiveness evaluations in at least 60 per cent of their protected areas.⁷

9. As mentioned above, the available information from the national reports, national biodiversity strategies and action plans and on protected areas suggests that countries are undertaking evaluations of

² For further information see document UNEP/CBD/COP/13/8/ADD2/REV1.

³ For further information see document UNEP/CBD/COP/13/8/ADD1/REV1

⁴ In preparation for the Ad Hoc Technical Expert Group on Indicators for the Strategic Plan for Biodiversity, convened in Geneva, Switzerland, from 14 - 17 September 2015

⁵ For further information see document UNEP/CBD/SBSTTA/19/4

⁶ See paragraphs 13 to 33 of document UNEP/CBD/SBSTTA/19/4

⁷ For further information see document UNEP/CBD/COP/13/INF/17

effectiveness to some degree. However the information that is available does not allow for the identification of any common methodologies, best practices or lessons learned. More detailed information from Parties on this issue would likely greatly facilitate the inclusion of such information in the sixth national reports by affording Parties an opportunity to learn from the examples and lessons of others.

III. APPROACHES TO EVALUATING THE EFFECTIVENESS OF MEASURES TAKEN UNDER THE CONVENTION

10. Evaluating the effectiveness of a measure undertaken as part of the pursuit of an objective differs from evaluating progress towards the objective, such as an Aichi Biodiversity Target. Evaluations of progress assess the extent to which an objective has been met. Evaluations of effectiveness aim to determine the extent to which a particular measure has brought about the desired change. In considering causes and effects within a field of multiple, often mutually reinforcing and sometimes contradictory courses of action, effectiveness evaluations are an important element of evidence-based policy making.

11. Numerous methodologies and approaches to evaluating effectiveness exist and as such this note should not be considered as an exhaustive discussion on the issue. Further in the context of this note, “measure” is taken to refer to any formal action taken by a government in support of the country’s implementation of the Convention. Measures could include policies, legislation and/or programmes and initiatives. Effectiveness evaluations can be carried out on broad policies, such national strategies, plans or programmes for the conservation and sustainable use of biological diversity or on more specific activities such as a the establishment of a protected area, the implementation of a payments for ecosystem services scheme or an invasive alien species eradication programme.

12. There is a large amount of academic literature, handbooks and guidance related to assessing the effectiveness of actions. This field of study is generally referred to as policy or programme evaluation.⁸ Approaches to assessing effectiveness can be broadly classified into three categories:

(a) Experimental: Experimental impact evaluations make use of randomly assigned control and experimental groups⁹. In this method only the experimental group receives the measure and the control group is used for comparison.

(b) Quasi-experimental: Quasi-experimental methods replicate experimental methods however the control group is not randomly assigned. As such this method is used in situations where a true experimental design cannot be applied but where it is still possible to identify an experimental and control group. This allows for a comparison between two groups to be made however because the control group is not randomly assigned the comparison may have limitations.

(c) Non-experimental: This approach does not use a control group. Only the experimental group is considered.

13. Both experimental and quasi-experimental evaluation methods require the use of control groups. For this reason they tend to be difficult to apply to the evaluation of biodiversity measures, particularly in real world settings, where the identification of suitable control groups is difficult if not impossible. Further experimental and quasi-experimental evaluations are generally most practical when assessing the

⁸ For example see: United Nations Evaluation Group (2016). *Norms and Standards for Evaluation*. New York: UNEG; United Nations Development Programme (2009). *Handbook on Planning, Monitoring and Evaluating for Development Results*. New York: UNDP; Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2016. *Impact Evaluation in Practice, second edition*. Washington, DC: Inter-American Development Bank and World Bank; Chen, Huey-Tsyh. 2005. *Practical Program Evaluation: Assessing and improving planning, implementation, and effectiveness*. Sage Publishing, United States of America, Crabbé, A. and P. Leroy, *The Handbook of Environmental Policy Evaluation*, Earthscan, 2008

⁹ In experimental and quasi experimental assessments the term “group” can refer to people, populations of plants and animals, or environments. It is a general term for a classification of a collection of things having similar characteristics.

effects of specific on-the-ground activities. They are more difficult, if not impossible, to apply for larger or more overarching policy measures. As a result, for most of the measures being implemented by countries, non-experimental approaches to evaluation are likely to be the most realistic evaluation method.

14. There are a variety of specific evaluation methods or approaches that can be applied in the non-experimental context. Some of these specific methods can be applied simultaneously or combined as part of the same assessment. They can also be applied by those individuals directly responsible for designing and implementing the measure (self-evaluation), they can be undertaken by or in collaboration with the partners involved in implementing the measure and/or the measure's beneficiaries (joint evaluation) or they can be undertaken by outside evaluators that have not been directly involved or impacted by the measure. For example many countries make use of national auditing bodies to undertake outside evaluations. Examples of specific methods include:¹⁰

(a) Programme Theory Evaluation – In this approach the evaluator considers the assumptions underlying the measure, its objectives and the tools used to implement it. The evaluator then makes an assessment of the overall logic of the measure and draws conclusions about the causal relationships that exist. A recent example of a programme theory evaluation is the fitness check, undertaken by the European Commission, of the European Union Nature Legislation (Birds and Habitats Directives).¹¹ As part of this check information was collected, including from consultations with members of the European Union and the general public. The analysis sought to identify changes that could be attributed to the intervention of the Directives being examined.

(b) Case Study Evaluation – This method evaluates how and why a given measure has worked or not by looking at a specific real world situation. One or several case studies could be used. A limitation with this approach is that it can be difficult to extrapolate the results and draw overarching conclusions as the observation are limited to the case study considered. An example of an evaluation using this approach is the review of effectiveness of Ireland's strategic environmental assessment directive undertaken by Ireland's Environmental Protection Agency.¹² One element of this evaluation was a review of 26 case studies.

(c) Formative/Developmental Evaluation – Using this method an evaluator considers how a measure was implemented. The evaluator does not consider the outcome of the measure but rather looks at the differences in how the measure was designed and how it was actually implemented. The objective is to develop recommendations on how similar measures could be better implemented in the future. This is primarily a qualitative approach to evaluation.

(d) Before-and-after comparisons – Using this approach conditions are assessed before and after the intervention. This approach assumes that any observed changes are the result of the intervention taken and that the impacts of the measure are observable at the time the evaluation is undertaken. This method is also called pre-test/post-test. An example of this approach is the evaluation of payment for ecosystem services programme implemented in Uganda. In this programme landowners were paid not to cut forests on their property. Landowners were surveyed and GPS and satellite imagery was used to determine tree cover before and after the introduction of the programme. In this way the evaluators could determine what the effect of the programme had been. Further, in this specific example the evaluators

¹⁰ The methods highlighted here draws heavily from Crabbé, A. and P. Leroy, *The Handbook of Environmental Policy Evaluation*, Earthscan, 2008,

¹¹ For further information see http://ec.europa.eu/environment/nature/legislation/fitness_check/index_en.htm.

¹² For further information see <http://www.epa.ie/pubs/advice/ea/reviewofeffectivenessofseainireland-executivesummary.html>.

used a randomized control group and therefore used a true experimental evaluation design.¹³ A further example of this type of a before and after comparison evaluation is what has been undertaken as part of the Sharkwatch Programme in the Maldives. In 2009 a reef shark fishing ban was introduced in the Maldives. Between 2009 and 2013 shark surveys were undertaken to determine what effect the ban had on shark populations.¹⁴

(e) Actual-versus-planned comparison – With this approach the anticipated outcomes of the measure are determined prior to the measure being implemented. Following the implementation of the measure the anticipated outcomes are compared with the outcomes actually achieved.

(f) Economic evaluation – This approach considers the outcomes and costs of an intervention in monetary terms. Generally information on the cost of implementing a measure is readily available, while information on the benefits of a measure is more difficult to determine. Furthermore, economic evaluation should also attempt to account for transaction costs as well as the overall social cost and benefits associated with the measure. There are two main types of economic evaluation:

- i. *Cost-benefit analysis* – This approach considers the total benefit from a measure in relation to what it cost to implement.
- ii. *Cost effectiveness analysis* – This approach compares the costs of two or more measures in reaching the same objective

15. These different evaluation methodologies have strengths and weakness and the most appropriate methodology to apply will depend on the type of measure implemented, the information sought, the overall purpose of the evaluation and national circumstances. The measures also greatly vary in terms of their feasibility, costs, and the robustness of the information generated. Further, in some cases multiple methodologies can be applied as part of the same assessment. Generally effectiveness evaluations are thought of as occurring once a measure has been completely implemented. However evaluations can also be undertaken when a measure is being designed, when it is being implemented or once it has concluded. For example Finland has recently completed a mid-term evaluation of implementation of its national biodiversity strategy and action plan.¹⁵

16. Regardless of the assessment approached used, information on which to base the assessment will be required. Given this, thought should be given to collecting necessary information throughout the development and implementation of the measure, including documenting any limitations the information may have. Depending on the methodology used, information collected before and/or after the intervention may be required. Relevant types of information include survey data, case studies, documentation related to the measure, budget information and administrative data.

17. Evaluations can also make use of expert opinion, and feedback from stakeholders. Indicators are also often used to assess effectiveness. However while indicators provide insights on the ultimate effects of measures, they do not always provide information on the process of implementing a measures and on the factors and context that have affected the success of it. In order to fully understand the effects of measures taken the progress measured by indicators needs to be causally linked to individual measures. A common method for establishing causal relationships is with the use of counterfactuals. Counterfactuals are comparisons of situations either with or without exposure to a measure, or before and after the application of a measure. They serve to identify the effect of different variables that may influence the outcome of a measure.¹⁶

¹³For further information see <https://www.povertyactionlab.org/evaluation/testing-effectiveness-payments-ecosystem-services-enhance-conservation-uganda>

¹⁴ For further information see <http://www.mrc.gov.mv/dv/programmes-and-collaborations/shark-resources-management/>

¹⁵ For further information see <https://www.cbd.int/doc/world/fi/fi-nbsap-v3-p3-en.pdf>

¹⁶ Miteva, D. A., S. K. Pattanayak, and P. J. Ferraro 2012, Evaluation of Biodiversity Policy Instruments: What Works and What Doesn't?, Oxford Review of Economic Policy Vol. 28 (2), p. 71.

IV. CONCLUSIONS

18. Information from the national reports, the national biodiversity strategies and action plans and other processes under the Convention on Biological Diversity suggests that most countries are undertaking effectiveness evaluations in some form. However there is little detailed information on this issue. This makes it difficult for countries to learn from the experience of others. More detailed national information on this issue could greatly benefit Parties in the context of numerous processes under the Convention, including the preparation of the sixth national reports, the discussions on the review of national experiences in implementing the Convention, and the preparation of a follow-up to the Strategic Plan for Biodiversity 2011-2020.

19. The available literature on evaluating effectiveness suggests that there are multiple documented options and approaches for assessing the effectiveness of measures. These approaches have different strengths and weaknesses as well as require varying amount of resources to implement. Given the different approaches that exist, Parties can choose the most appropriate approach for assessing the effectiveness of measures taken in line with their NBSAPs. In fact many Parties, in their fifth national reports, indicate the need to undertake assessments of effectiveness or note that plans were already underway to do so.