Evaluation of the Belmont Forum: Final report

Judith Vermeer (project leader), Etienne Vignola-Gagné, David Campbell, Ivette Oomens, Henrique Pinheiro, Liana Petrosova, Derek Eaton, Geert van der Veen

www.technopolis-group.com

Evaluation of the Belmont Forum, Final report

Version 4-9-2020

Judith Vermeer (project leader), Etienne Vignola-Gagné, David Campbell, Ivette Oomens, Henrique Pinheiro, Liana Petrosova, Derek Eaton, Geert van der Veen

Table of Contents

Tables

Figures

Abbreviations and acronyms

AMI	Altmetrics mentions index
ANR	French National Agency for Research
ARC	Average of relative citations
BF	Belmont Forum
BFgo	Belmont Forum Grants Operations
BMBF	German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung)
CDC	Citation distribution chart
CDI	Citation distribution index
CEH	Climate, Environment and Health
CONACYT	National Council of Science and Technology (Mexico)
CRA	Collaborative Research Action
CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
Ctries	Countries
DR	Disaster Risk, Reduction and Resilience
EC	European Commission
ERA	European Research Area
ERC	European Research Council
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
FTE	Full-time equivalent
GCR	Global change research
GPC	Group of program coordinators
HAP10%	Share of highly altmetrics-mentioned publications
HAP1%	Share of highly altmetrics-mentioned publications
HCP10%	Share of highly cited publications
HCP1%	Share of highly cited publications
HIP10%	Share of highly interdisciplinary publications
HLPF	High-Level Political Forum
HMP10%	Share of highly multidisciplinary publications
HR	Human resources
IAI	Inter-American Institute for Global Change Research

ICR	International co-publication rate
IGFA	International Group of Funding Agencies
IGO	Intergovernmental organization
IIASA	International Institute for Applied Systems Analysis
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
ISC	International Science Council
ISR	Inter-sectoral co-publication rate
MESRI	French Ministry of Higher Education, Research and Innovation
MI	Multidisciplinarity index
MINCyT	Ministry of Science, Technology and Productive Innovation (Argentina)
MoU	Memorandum of understanding
NGO	Non-governmental organization
NERC	UK Natural Environment Research Council
NSERC	Natural Sciences and Engineering Research Council of Canada
NSF	US National Science Foundation
NSFC	National Natural Science Foundation of China
OA	Open access
ORC	Other research centers
SC	Belmont Forum Steering Committee
SDG	Sustainable Development Goals
SEI	Science-driven e-infrastructure innovation for the enhancement of transnational, interdisciplinary and transdisciplinary data use in environmental change
T2S	Transformations to Sustainability
ToR	Terms of reference
TPO	Thematic Programme Office
TSRI	Thailand Science Research and Innovation
UN	United Nations
UNEP	United Nations Environment Programme
WCS	Weighted CiteScore

Executive summary

This report presents the results of the external evaluation of the Belmont Forum that was executed in the period September 2019 to August 2020 by Technopolis Group and Science-Metrix.

The Belmont Forum is a partnership of funding organizations, intergovernmental organizations, international science councils, and regional consortia committed to the Belmont Challenge:

> *"International transdisciplinary research providing knowledge for understanding, mitigating and adapting to global environmental change."*

The Forum was established in 2009 and started in its present form in 2014. At this point, more than €150m of funding was awarded to 132 projects, funding more than 2,000 scientists around the globe (covering 73 countries).

The Forum's scientific impact is impressive. The Belmont Forum made significant contributions to the scientific communities working on environmental change and associated topics. The Forum scored well ahead of several of its benchmarking groups and achieved similar results as that of the ERC. The Belmont Forum also had a significant effect on policy. In terms of networking effects, the Belmont Forum has been effective in generating new collaborations and partnerships across actors, disciplines, and countries. Not only did the Forum have a very strong networking effect between funding organizations, it is also successful in fostering networking between researchers of different disciplines and countries. Despite the high concentration of countries from the Global North as members, the Forum did in fact manage to establish collaboration between the Global North and the Global South. In the co-publication analysis, the Belmont Forum scored significantly higher in terms of North–South collaboration than other funders. However, there is still room for improving the involvement of Southern researchers and organizations into the Forum.

The Belmont Forum's vision, objectives, and activities are still highly relevant today. The Forum's governance structure is practical and functional, the policies are coherent and the operations efficient.

To remain sustainable, the understaffing of the Secretariat and an increasing divergence among the members in terms of participation and commitment need to be addressed.

Various recommendations are presented.

Extended summary

This report presents the results of the external evaluation of the Belmont Forum that was executed in the period September 2019 to August 2020 by Technopolis Group and Science-Metrix. The evaluation assessed impact and organization. The evaluation covers the period 2009 (when the Belmont Forum was established) until mid-2020. For the organizational evaluation, the focus is on the period 2017–2020 (suitable information available and most relevant). It should be noted that in some instances, data was only available for more restricted time periods (as indicated in the relevant tables).

The evaluation consisted of a combination of quantitative and qualitative data collection and analysis tools: desk study and data analysis, 40 interviews, an online survey of beneficiaries, an online survey of members and partners, two case studies, and bibliometric, altmetric and social network analysis.

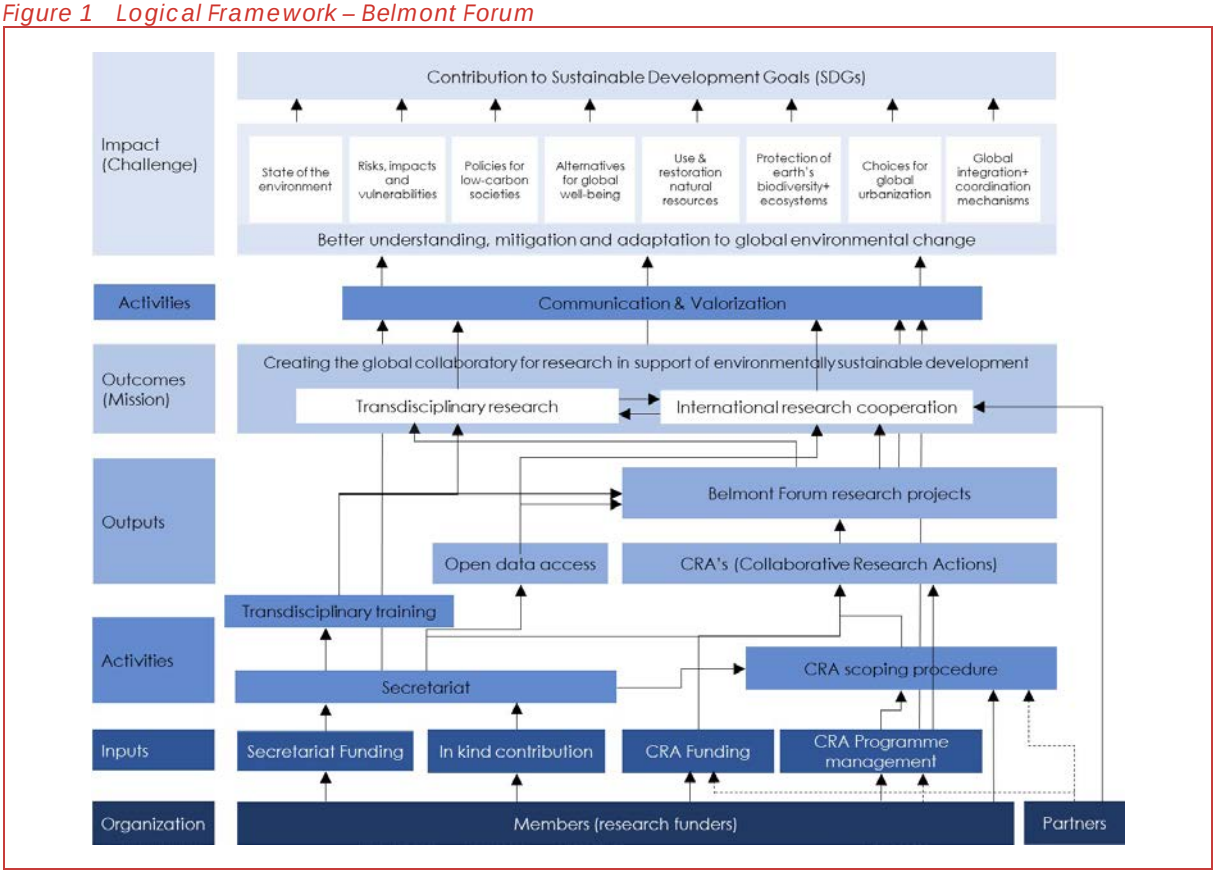
Background, Belmont Forum goals

The Belmont Forum was established in 2009 and started in its present form in 2014. It is a partnership of funding organizations, intergovernmental organizations, international science councils, and regional consortia committed to the Belmont Challenge (2016 formulation):

> *"International transdisciplinary research providing knowledge for understanding, mitigating and adapting to global environmental change."*

Figure 1 below presents a Logical Framework for the Belmont Forum.

Figure 1 Logical Framework – Belmont Forum



Source: Technopolis Group and Science-Metrix (2020)

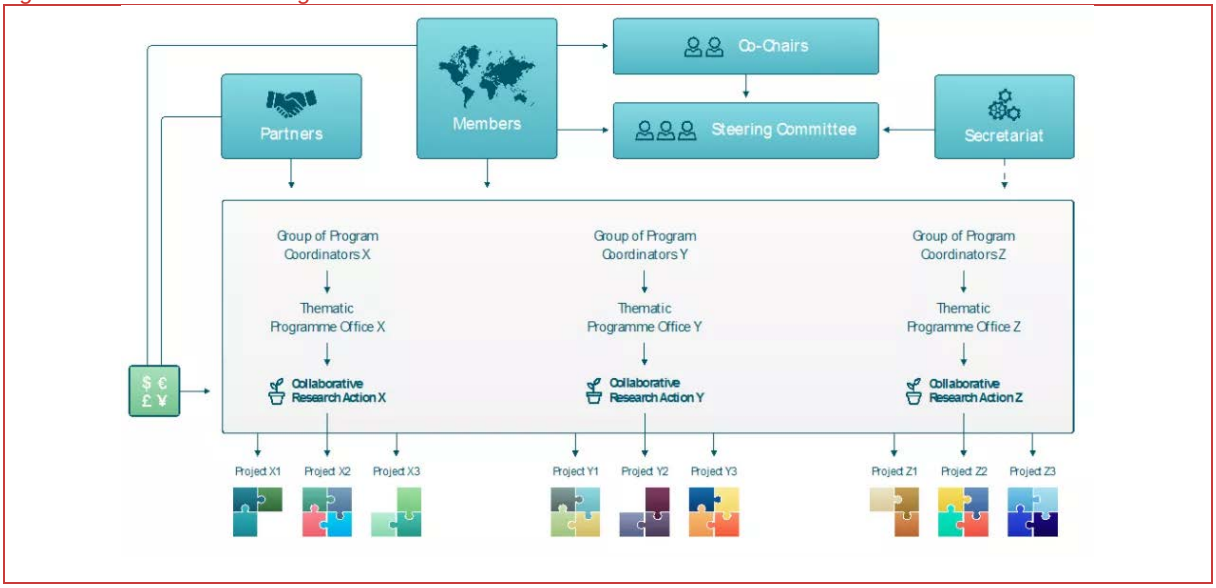
Belmont Forum activities

The Belmont Forum funds research projects that contribute to the Belmont Challenge under Collaborative Research Actions (CRAs). Under each CRA, the Belmont Forum institutes international calls for proposals in partnership with interested organizations. Sixteen CRAs have been developed by the Belmont Forum, from 2012 until today. Not all Belmont Forum members and partners participate in every CRA, but every CRA should at least include three Belmont Forum members. At this point, more than €150m of funding was awarded to 132 projects, funding more than 2,000 scientists around the globe (covering 73 countries).

Organization and governance

The permanent governance structure of the Belmont Forum consists of members, partners, a Steering Committee, co-chairs, and the Secretariat (Figure 2).

Figure 2 Belmont Forum governance



Source: Belmont Forum website (2019)

Members meet at least once every year in the Plenary Meeting, which is the highest decision-making level of the Forum. At present (July 2020) the Belmont Forum has 30 members and six partners. The *Steering Committee* (SC) is elected by the Belmont Forum members to support and advise the Belmont Forum activities between plenary meetings. The committee currently consists of eight members headed by two co-chairs. The Belmont Forum *Secretariat* is the executive arm of the Forum and is tasked with carrying out decisions of Plenary Meetings and SC. *Partners* are organizations that subscribe to the Belmont Challenge, but do not have a vocation to meet the criteria for membership. This overall governance structure of the Belmont Forum is overseeing and supporting specific governance bodies at the level of each CRA: consisting of *Thematic Program Offices* (TPO) entrusted with the preparation, publishing and management activities during the period of a CRA implementation and *Group(s) of Program Coordinators* (GPC)**,** responsible for the practical implementation of a given CRA. GPCs are open to any Belmont Forum member interested in the CRA but also to other funding organizations, which may join a CRA without becoming a Belmont Forum member.

Relevance

The Belmont Forum's vision, objectives and activities are still highly relevant today. The rationale for the establishment of the Belmont Forum is highly recognized at international political levels and the issue of global environmental change as well as the collaborative transdisciplinary approach of the Forum, which is at the core of the Belmont Forum action, have remained extremely important. While there are many forums and organizations that focus on environmental change, few of them offer active multilateral funding programs that specifically spur global transdisciplinary research.

Vis-à-vis its members (which are mostly funding organizations), the Belmont Forum serves two relevant purposes – namely, (1) engaging a dialogue with the global research community on environmental challenges (discussion and potential alignment of research agenda), and (2) strengthening the transdisciplinary approach needed to address the complexity and global character of the problems. Both are key motivations to join the Forum.

However, after 10 years in operation, the Belmont Forum is entering into a maturity stage. The increasing number and diversity of its members and the need for achieving a critical mass of members to support its operations, suggest that there is a need to put the strategic direction of the Forum on the agenda and to discuss more explicitly what the Forum wants to achieve in the future.

Effectiveness and impact

The Belmont Forum's scientific impact is impressive. In the bibliometric analysis, the Forum scored well ahead of several of the benchmarking groups and achieved similar results as that of the ERC (known for its excellent research). The Belmont Forum made significant contributions to the scientific communities working on environmental change and associated topics, even above and beyond what might have been expected.

The Belmont Forum also had a significant effect on policy, particularly scientific advice and evidence synthesis aimed at policymakers. The evaluation findings show that scientific publications funded by the Forum were cited quite often in policy documents. Almost a third of the publications funded in the period 2012–2014 were cited in one or more policy documents (e.g. from the EU, FAO, UNEP, and the World Bank). The frequent policy citations of BF-funded publications can be considered a reliable indicator that the Forum has provided meaningful input for science advice and argumentation by a range of national governmental agencies, IGOs and think tanks.

In addition to the Forum's impact on the scientific communities and policymakers, it seems Belmont Forum has also been relatively effective in disseminating knowledge and fostering discussion in wider circles and audiences. Forum publications scored well above the average world level for their OA accessibility (although lower than some benchmarking groups). BF-funded publications have been taken up and discussed by online and social media communities to a good degree, depending on the exact dimension considered. Forum publications performed best in the benchmark when it came to mentions in journalistic news items and on Facebook, did slightly less well on Twitter and were surpassed by a small group of benchmarking groups for Wikipedia citations. Turning to web citations such as blog posts, videos, policy briefs or research tools and data sets, the main conclusion is that impact appears to have been achieved, but that definitive evidence of these achievements could not yet be collected.

In terms of networking effects, the Belmont Forum has been effective in generating new collaborations and partnerships across actors, disciplines, and countries. Not only did the Forum have a very strong networking effect between funding organizations, it is also successful in fostering networking between researchers of different disciplines and countries.

All in all, the Belmont Forum had considerable positive effects, which are likely to contribute to enhancing our understanding of global environmental change, as well as our ability to mitigate and adapt to this change.

Efficiency and organizational set-up

The Forum's governance structure is practical and functional, it supports the achievement of the objectives, and it divides control and responsibilities in a clear and balanced way.

With the adoption of the memorandum of understanding (MoU) the member organizations agreed on launching calls on a periodic basis using a model of a joint call and evaluation but coordinated funding, according to organizational rules, with no requirement for funding across borders (but this can be done when their organizational mandate makes this possible, in this case it often allows more inclusive participation, particularly from low and middle income countries. It also describes how members of the Forum interact in the calls. This part of the governance works well and leads to impact. A large majority of the surveyed beneficiaries are satisfied or very satisfied with the Forum's processes around the CRAs.

Where the focus of the MoU is on the calls, the terms of reference (ToR) describe (predominantly) the organization and procedures of the Forum outside the calls. Here the role of the Secretariat is very important. It provides support to the SC and the co-chairs and organizes the Plenary Meeting and provides support to the CRAs in many ways: in the scoping phase, with ICT (BFgo: helping with the launch, answering questions submitted by proposers), training (e.g. training of researchers in transdisciplinary research) and in monitoring the project progress and implementation (e. g. organizing progress workshops and valorization meetings).

Total costs for the Secretariat (excluding costs for this external evaluation, but including all activities of the Secretariat, including SC and plenary support and communication), and including the costs for the two AAAS fellows and BFgo costs (financed by NSF) and IAI contribution, bring the costs at around 4%–5% of CRA budget. In the experience of Technopolis, program management costs for straightforward national programs are in the range of 7%–10%. Higher percentages (12%–15%) are not unusual for complex programs requiring much coordination. The Belmont Forum is such a complex program, and the Secretariat is managing quite a bit of the complexity. Although a fair amount of CRA preparation costs and the full costs for project selection and financial administration are with the members (and could not be provided when the evaluators asked for this), the ratio of execution costs against program funding seems to be low for the Belmont Forum. In this respect Belmont Forum Secretariat seems to be operating efficiently.

The present composition of the Secretariat with one director, two full-time AAAS fellows (an in-kind contribution by the NSF) and four part-time (0.2 FTE each) staff (as in-kind contributions of other members) is, however, not sustainable. The staff is overcharged, and despite the high effectiveness and efficiency as presented above, there is increasing dissatisfaction with (some) members (especially about communication).

The rules and procedures of the Belmont Forum are transparent and inclusive. Despite the high concentration of countries from the Global North as members, the Forum did in fact manage to establish collaboration between the Global North and the Global South. In the co-

publication analysis the Belmont Forum scored significantly higher in terms of North–South collaboration than other funders (including NSFC, BiodivERsA, and the ERC) – for example, when looking at the share of papers containing at least one author from an OECD ODA country and at least one from a non-ODA country. However, there is still room for improving the involvement of Southern researchers and organizations into the Forum.

The internal CRA procedures (i.e. scoping of the calls, drafting of proposals, selection of the research proposals) are satisfactory to the members and the beneficiaries. The Forum seems to be an efficient mechanism in terms of coordinating and promoting the proposal processes.

Coherence

The Belmont Forum's thematic prioritization through the CRAs as well as the projects funded by the Forum are coherent with each other. The CRAs (and related projects) are well-chosen and each have a clear link to the Forum's overall mission and objectives. While some CRAs are quite closely related or sometimes even overlap, we do not consider this as problematic for the overall coherence of the Forum. This may help to understand certain subject matters in more depth, and to look at similar issues from different angles.

Regarding the alignment of the Forum with other international programs, the Forum is quite a unique initiative being fully international, focused on transdisciplinary research, and global environmental change. However, there are some overlaps with existing programs or organizations as reported by the members. There is a need for increased external engagement with funding and donors' organizations (aid agencies, philanthropic foundations, etc.) and communication, particularly in the scoping phase of the CRAs.

Added value

The Belmont Forum offers clear added value within the global landscape of science funding and a unique combination of project-level strengths and achievements. It fosters multilateral collaboration among funding agencies as well as among researchers (including North–South collaboration).

For researchers, it provides access to research funding, and particularly to support transdisciplinary research projects. The bibliometric analysis shows that the Belmont Forum is worldwide leading in terms of interdisciplinarity and multidisciplinarity. Bibliometric indicators for transdisciplinarity are still being developed. Belmont Forum is working on such indicators, together with other supporters of transdisciplinary research.

Turning to indicators of broad dissemination and societal uptake of findings, Belmont Forum's publications displayed a high level of impact within science advice and governmental research documents. The Forum was trailed closely only by the ERC.

Sustainability

It was not possible and not in the scope of this evaluation to conduct an in-depth review of the research projects and their results that would have permitted to precisely appraise the sustainability of their results. However, members and partners of the Belmont Forum provided anecdotal evidence of sustainability. Researchers, in a large portion, have disseminated their project results by combining traditional channels (scientific publications and conferences) with discussions with policymakers (60% of respondents) and non-scientific publications (47%). The

fact that the dissemination of project results to various target audiences was largely effective is a promising result in terms of sustainability.

Regarding the sustainability of the Forum, as an organization, there are two challenges to consider and to address in a relatively short term. First, the most pressing issue is the staffing and financing of the Secretariat. Second, there is an increasing divergence among the members in terms of participation and commitment.

Recommendations

Recommendation 1: This evaluation shows that the Belmont Forum is relevant and achieving (very) good results in an efficient manner. The Belmont Forum should consider the results of this evaluation as an encouragement to build on what is already achieved and continue to initiate and support international transdisciplinary research providing knowledge for understanding, mitigating and adapting to global environmental change.

Recommendation 2: The growth in members and activities means that the Belmont Forum is entering into a maturity stage. At this point there is a need to discuss whether all members are still in agreement with the Belmont Forum's mission and ensure a common agreement of the appropriate way forward. The following questions should be considered to (re)create strategic alignment of members:

- What objectives should the Forum pursue? Should there be more focus, or less?
- What more can be done to include more member organizations and more researchers from the Global South?
- What more can be done to encourage the use of research results from the CRAs?

The common agreement resulting from the discussion should be operationalized and reflected in the Forum's mission, objectives, and in the formulation of CRAs.

Recommendation 3: The growth in members and activities the demands on the Secretariat have also increased, and the Secretariat is understaffed to meet up to the ambitions/expectations. We suggest increasing the capacity of the Secretariat to 5–6 FTE staff members.

Recommendation 4: To bring the Secretariat at the level of 5–6 FTE, around €800,000 per year (in cash or in kind) would be necessary to cover costs. To afford this increase of membership fees and/or increase in number of members (or other financial contributors) are necessary. Differentiation of contribution (higher fees for "richer" organizations, or fees depending on the participation in/contribution to CRAs) might be an option. Also increasing in-kind contributions to, for example, 0.4 FTE/organization is possible. Finding new members/contributors is not easy and will require large efforts from the Secretariat (or Members) to realize. The alternative however is a serious reduction of tasks for the Secretariat, which would either mean a serious reduction in the ambition levels of the Belmont Forum and/or a transfer of tasks presently done by the Secretariat to the member organizations.

Recommendation 5: In all scenarios for solving present understaffing of the Secretariat, more explicit prioritization of activities of the Secretariat is needed. We suggest addressing this with a more explicit annual planning cycle, with short, to the point, annual work plans with clear priorities (so that activities of the Secretariat match capacity), annual reporting to the Plenary Meeting, empowering of the Secretariat director to implement these annual plans (including more explicit budget responsibility), financial and progress reporting in every SC meeting, and an active interaction between SC/co-chairs and Secretariat. The more explicit steering should

also include an explicit annual assessment of the performance of the Secretariat Director by the co-chairs and annual accounts approved by an external accountant.

***Recommendation 6*:** The impact of the Belmont Forum increases when results of Belmont Forum projects are applied broader than in the project setting itself. Although there is already quite some attention for result communication of research results (and this evaluation shows there is also quite uptake of these results in policy circles), there seems room to further boost the online visibility and uptake of both Forum publications and non-journal outputs. In case the Belmont Forum decides to increase capacity at the Secretariat, and a dedicated communications officer is appointed, s/he may have this is one of her/his tasks. Developing "outcomes narratives" for each project and presenting them in annual reports, valorization reports, on a dedicated section of the Belmont Forum website, and/or on specialized platforms such as Kudos or Researchfish would increase exposure and consequently impact.

1 This evaluation

1.1 This evaluation

This report presents the results of the external evaluation of the Belmont Forum (hereafter "the Forum") that was commissioned by the Forum and executed in the period September 2019 to August 2020 by Technopolis Group and Science-Metrix. The main objective of the evaluation was: *"to evaluate its progress towards meeting the Belmont Challenge and the efficacy of the organization to continue to reach its goals and fulfil its mission."*[1]

The evaluation consisted of two main components:

- **Impact evaluation:** an assessment of the delivery of the CRAs, including their scientific, policy, and other impacts and the added value of transnational and transdisciplinary collaborations.
- **Organizational evaluation:** an exploration of the effectiveness of the Forum (including the funding mechanisms, its governance and management (plenary, Steering Committee and Secretariat) efficiency of procedures, transparency, inclusiveness, etc.).

The evaluation covers the period from 2009 (when the Forum was established) until end-2019. For the organizational evaluation, the focus is on the period 2017–2019 (suitable information available and most relevant, as the aim of the evaluation is primarily to improve Forum's performance further).

1.2 Evaluation questions

Based on the ToR for the evaluation, we defined 17 evaluation questions (Table 1).

Table 1 Overview of evaluation questions

#	Evaluation question
Relevance	
1	How relevant are the **mission and objectives** of the Forum (in the context of understanding, mitigating, and adapting to global environmental change, as well as achieving the SDGs)?
Effectiveness and impact	
2	How successful has the Forum been in effectively **engaging with its key (target) stakeholders**?
3	**Networking effects:** How effective has the Forum been in generating new collaborations and partnerships across various sectors, disciplines, and countries around the globe? To what extent are the funded projects truly co-designed and co-created?
4	**Scientific outcomes:** To what extent did the Forum contribute to the science base for environmental change (understanding, mitigation, and adaptation)?
5	**Wider dissemination of knowledge:** How effective has the Forum been in disseminating knowledge and other outputs generated by the Forum? To what extent were results of the Forum disseminated, taken up and discussed beyond academic circles?
6	**Policy effects and outcomes:** To what extent did results of the Forum foster policy debate or developments at international and national level or facilitate policymaking / implementation?

[1] Belmont Forum (2019). Request for Proposals for the Evaluation of the Belmont Forum

#	Evaluation question
7	**Other effects and outcomes:** Did the Forum have any other effects (intended or unintended, positive or negative)?
8	**Overall impact:** To what extent has the Forum (in collaboration with all its stakeholders) contributed to the Challenge of understanding, mitigating, and adapting to global environmental change?
Efficiency and organizational set-up	
9	How appropriate are the **governance and organizational set-up** of the Forum (especially considering its growing membership)?
10	Are the **rules and procedures** of the Forum appropriate (e.g. transparent and inclusive) and are they regularly reviewed for appropriateness?
11	To what extent has the Forum **carried out its work efficiently** (e.g. financial and human resources, internal procedures)?
12	Are the Forum's **financial and human resources** appropriate for the work entrusted to it?
Coherence	
13	**Internal coherence:** Are the CRAs and funded projects coherent with each other?
14	**External coherence:** Are the activities of the Forum coherent with other initiatives in the context of environmental change?
Added value	
15	What is the **added value** of the Forum (compared to other initiatives at various governance levels)?
Sustainability	
16	To what extent are the **effects and results** of (projects funded by) the Forum sustainable in the future?
17	To what extent is the Forum **sustainable in the future**?

Source: Technopolis Group and Science-Metrix (2020)

1.3 Evaluation methodology

The evaluation consisted of a combination of quantitative and qualitative data collection and analysis tools: desk study and data analysis, 40 interviews, an online survey of beneficiaries, an online survey of members and partners, two case studies, and bibliometric, altmetric and social network analysis. More information about the methodology can be found in the appendices.

1.4 This report

The report is structured as follows:

- Chapter 2 presents a brief introduction to the Forum.
- Chapter 3 presents the main findings of the evaluation structured around the main evaluation criteria of relevance, effectiveness and impact, efficiency and organizational set-up, coherence, added value, and sustainability.
- Chapter 4 presents overall conclusions and recommendations with suggestions for future improvements.

The technical appendices (separate document) contain more detailed study findings, structured by data collection method.

- Appendix A provides an overall description of the methodology including paragraphs on the difficulties of measuring impacts of transdisciplinary research.

- Appendix B presents the methodology and findings of the survey of beneficiaries of the Forum.
- Appendix C presents methodology and findings of the survey of Forum members and partners.
- Appendix D contains a list of interviewees that we spoke to, and the detailed interview guides that were used for the interviews.
- Appendix E describes the findings of the two case studies conducted for this evaluation.
- Appendix F contains the full technical report of the bibliometric and altmetric analysis.

2 The Belmont Forum

2.1 Establishment and rationale

The Forum is a partnership of funding organizations, intergovernmental organizations, international science councils and regional consortia committed to the advancement of transdisciplinary science in the field of global change. The origins of the Forum date back to 1990 with the International Group of Funding Agencies (IGFA) that formed as an informal group to support global change research (GCR).[2] In 2009, the Forum was established as a council of principals for the IGFA. In 2014, the two organizations were merged under the name Belmont Forum.

The Forum is based on the understanding that, to address the challenges of global environmental change in the 21st century, **societies need to be informed by high-quality research.** The Forum feels that despite the currently existing knowledge about how humans transform the global environment and how these changes may affect human well-being, it is necessary to further improve understanding of not only the impacts, vulnerabilities and risks, but also the opportunities of environmental change in order to really be able to address environmental change. By a **transdisciplinary** approach, across scientific disciplines and co-developing and co-implementing research with stakeholders, better adaptation and mitigation strategies can be developed and applied, and society can better benefit from opportunities in a sustainable manner. Such knowledge should enable effective decision-making and support equitable economic and social development.[3] The Forum supports multi-national and transdisciplinary research, bringing together natural sciences, social sciences, and the humanities as well as stakeholders in co-creating the knowledge and solutions for sustainable development that benefit society. Funding agencies and international scientific associations from six continents are either members or partners.

Recognizing the importance of the United Nations **Sustainable Development Goals** (SDGs) as a framework for sustainable development at global level, the Forum wants to make a considerable contribution to the implementation of the SDGs under existing conditions of global change by supporting relevant interdisciplinary research.[4]

2.2 Goals and objectives

The members and partners of the Forum are committed to fulfill the **Belmont Challenge**:

> *"International transdisciplinary research providing knowledge for understanding, mitigating and adapting to global environmental change."*[5]

To address the Challenge, there is a perceived need to change the way global environmental change research is supported and undertaken. The Forum does this from a research funders' point of view, where international resources for research are coordinated and scoped in a cross-community framework.

---

[2] Belmont Forum (2019). History. Retrieved on 17.06.2020 from: http://www.belmontForum.org/about/#History

[3] Belmont Forum (2019). History. Retrieved on 17.06.2020 from: http://www.belmontForum.org/about/#History

[4] Belmont Forum (2017). Belmont Challenge White Paper (November 2017, revised in April 2017)

[5] Belmont Forum (2019). History. Retrieved on 20.07.2020 from: https://www.belmontforum.org/about/

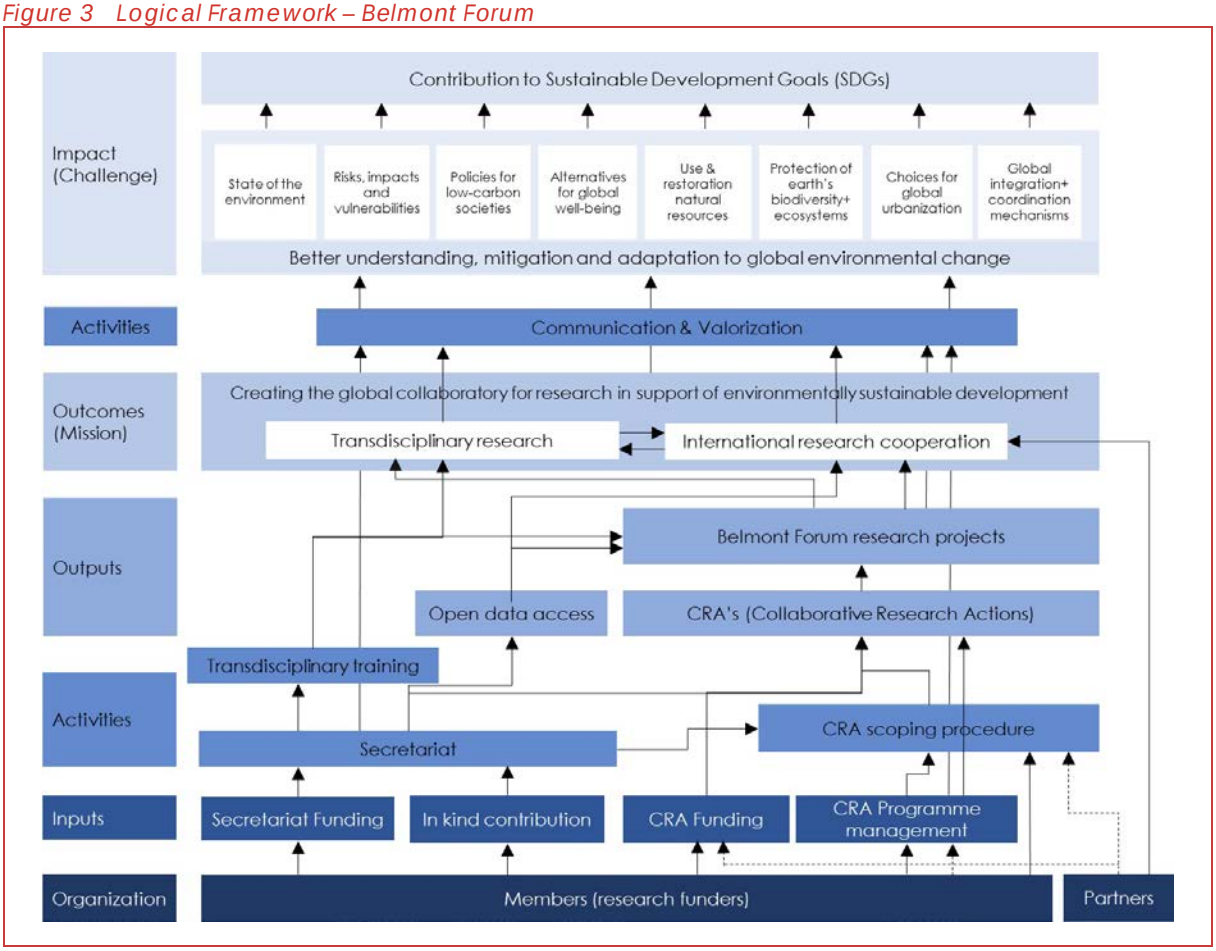
The **mission** of the Forum could therefore be formulated as "*creating the global collaboratory for research in support of environmentally sustainable development.*"[6] More specifically, the Forum supports inter- and transdisciplinary research which takes account of coupled natural, social, and economic systems, to promote and enhance:

- *"Information on the state of the environment, through advanced observing systems, and enhanced environmental information service providers to users;*
- *Assessments of risks, impacts and vulnerabilities, through regional and decadal analysis and prediction;*
- *Evaluation of policies that lead to low-carbon societies and consider how best to implement international and national commitments on emissions reductions and sustainable development;*
- *Analysis of alternatives that promote global well-being, considering the different needs of developing and developed economies;*
- *Studies on how best to use and restore our natural resources on land, water and energy in a sustainable and efficient way, considering global teleconnections, and focusing on sustainable production of goods for our societies while mainstreaming strong environmental protection;*
- *Protection of Earth's biodiversity and endangered ecosystems;*
- *Examination of choices for managing global urbanization and pathways towards more sustainable cities;*
- *Analysis of global integration and coordination mechanisms, to address interdependencies and marshal the necessary resources."*[7]

[6] Belmont Forum (2015). Belmont Forum Terms of Reference

[7] Belmont Forum (2017). Belmont Challenge White Paper (November 2017, revised in April 2017)

Figure 3 Logical Framework – Belmont Forum



Source: Technopolis Group and Science-Metrix (2020)

Since this 2016 update to the white paper, the Forum has broadened activity so that knowledge relevant to the Challenge is engaged through the lens of individual CRAs, now including health, information sciences, and so forth.

Based on the Forum's mission and goals, Figure 3 above presents a **Logical Framework for the Forum**. It presents the variety of actors involved in the Forum, the dynamics that they aim to bring in motion, and the long-term objective of ultimately contributing to a better understanding, mitigation and adaptation to global environmental change, in its turn contributing to achieving the SDGs.

2.3 Belmont Forum activities

The Forum funds research projects that contribute to the Belmont Challenge under **CRAs**. Under each CRA, the Forum institutes international calls for proposals in partnership with interested organizations. Not all Forum members and partners participate in every CRA, but every CRA should at least include three Forum members. Sixteen CRAs have been developed by the Forum, from 2012 until today, listed in Table 2. Over the years, more than 136 funded projects with a total award amount of more than €154 million.

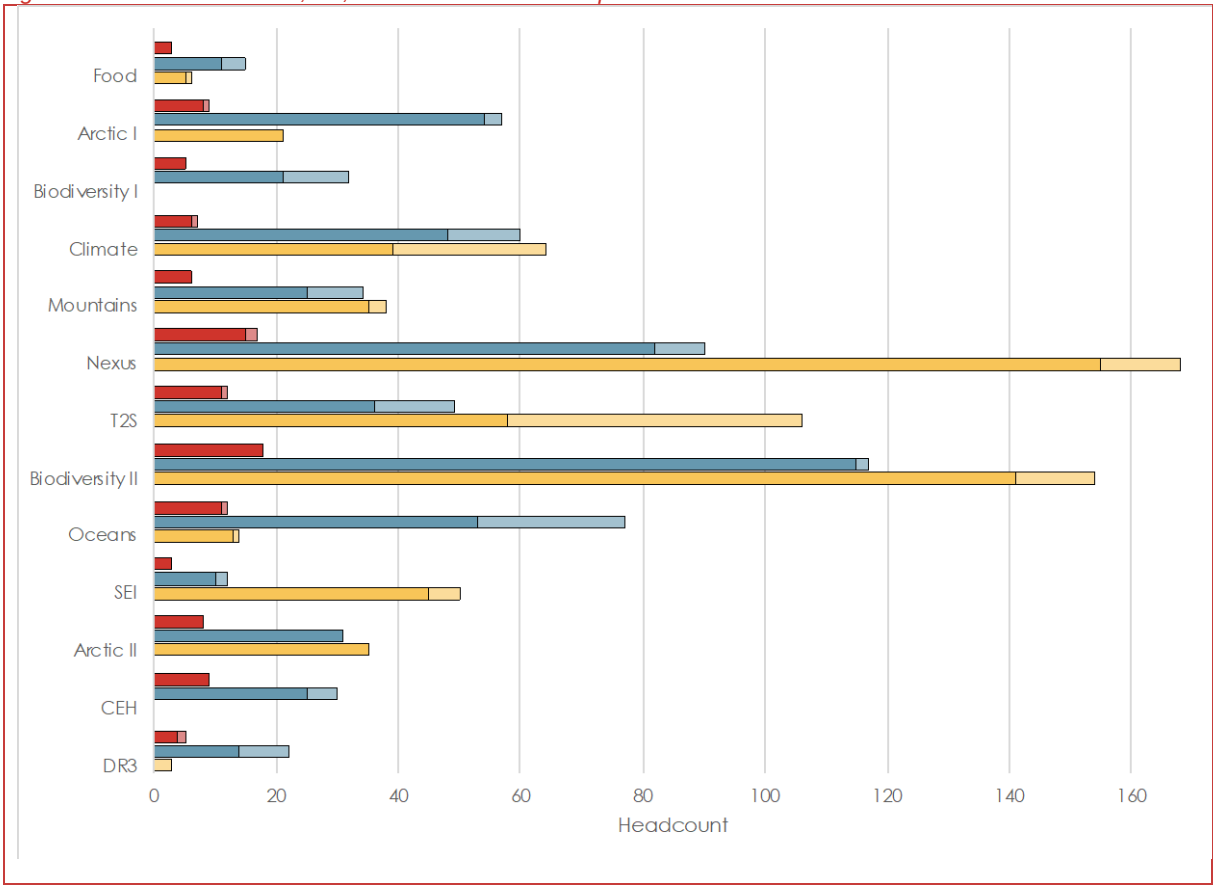
Table 2 CRA Success rates

CRA	Year	Registrations	Full Applications	Projects/ Awards	Application Success[1]	Project Funding Agencies	Project Personnel	Project Countries	Total Amount Awarded (€)
Coastal Vulnerability	2012	61	9	7	78%	11	85	21	10.781.968,10
Freshwater Security	2012	76	9	6	67%	10	57	11	7.793.408,54
Food Security and Land Use Change	2013	24	17	7	41%	3	63	12	5.670.919,87
Arctic I: Arctic Observing and Research for Sustainability	2014	N/A	46	10	22%	5	89	12	7.767.350,00
Biodiversity I: Scenarios of Biodiversity and Ecosystem Services	2014	N/A	9	4	44%	4	69	10	1.333.893,00
Climate Predictability and Inter-Regional Linkages	2015	39	16	8	50%	5	287	18	13.056.241,00
Mountains as Sentinels of Change	2015	83	27	6	22%	4	143	11	5.884.637,65
Data: E-infrastructures and data management	2015	Data not available							
Nexus: Sustainable Urbanization Global Initiative / Food-Water-Energy	2016	76	39	16	41%	10	448	19	30.157.752,88
T2S: Transformations 2 Sustainability	2016	274	39	12	31%	6	263	29	11.659.133,17
Biodiversity II: Scenarios of Biodiversity and Ecosystem Services	2017	143	135	22	16%	10	480	25	27.576.803,00
Transdisciplinary Research for Ocean Sustainability	2018	58	22	13	59%	5	108	21	9.435.827,70
SEI: Science-driven e-infrastructure innovation for the enhancement of transnational, interdisciplinary	2018	N/A	9	3	33%	2	83	7	3.274.347,59
Arctic II: Resilience in Rapidly Changing Arctic Systems	2019	N/A	24	8	33%	6	86	10	4.592.079,70
Climate, Environment and Health	2019	78	59	9	15%	0	69	20	9.730.908,80
DR3: Disaster Risk, Reduction and Resilience	2019	N/A	34	5	15%	1	44	18	5.389.963,13
Total		912	494	136	28%	54	2374	102	154.105.234,13

[1]Ratio between awards to full applications; Please note that the data for the most recent CRAs is incomplete at the time of writing.

The figure below shows the number of consortium leads, PIs, and team members per CRA. In doing so, it distinguishes between those from the Global North and those from the Global South.

Figure 4 Consortium leads, PIs, and Team Members per CRA



This figure shows Consortium Leads (●/●), PIs (●/●) and Team Members(●/●) per CRA. The left part of the bar (brighter colour) representing participants from the Global North, while the right side of the part depicts the number of participants from the Global South.

2.4 Organization and governance

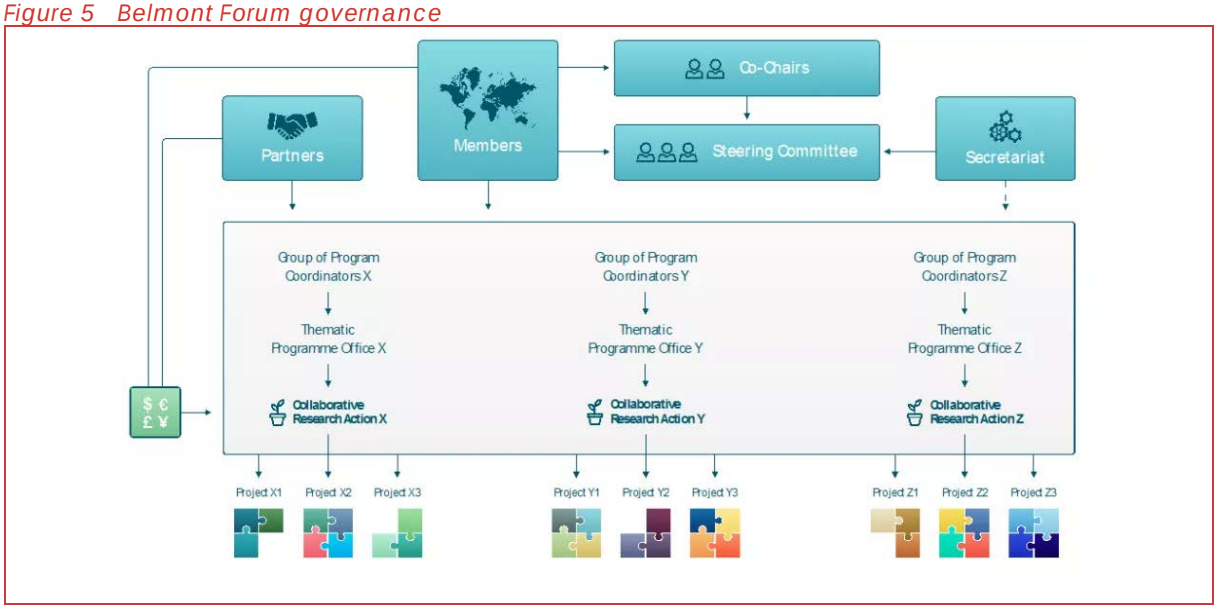
The Forum does not constitute a formal legal entity, but is an informal high-level group of funders of global environmental change research and international science partners/organizations and it is a member-based organization, fully focused on organizing and delivering international collaborative calls for research proposals and valorizing results from those calls. The rules for these calls are determined by the **memorandum of understanding** (MoU) that was signed between the existing members (formally: Forum Partner Organizations) in 2012, and that is agreed by all new members and those contributing funding or in-kind resources to the Forum research as well. The rules for the organization of the Forum outside the CRAs are determined by the **terms of reference** (ToR) for the Forum (ToR last revised 2015, renewed revision discussed in Taipei, 2019).

The permanent governance structure of the Forum consists of **members**, **observers**, **partners**, a **Steering Committee (SC)**, **co-chairs**, and the **Secretariat**. Members and partners cover six continents.[8] Each (CRA) has its own governance, with a **Thematic Programme Office** (TPO)

[8] Belmont Forum (2017). Belmont Challenge White Paper (November 2017, revised in April 2017)

and a **Group of Program Coordinators** (GPC), as well as an implementation plan that is agreed upon by the GPC but ultimately is in accordance with the ToR and MoU (Figure 5).

Figure 5 Belmont Forum governance



Source: Belmont Forum Website (2019)

Members have (1) formally declared their commitment and willingness to actively contribute to Forum operations (with their work force, financial resources, and existing research and innovation investments, etc.), (2) have participated in at least one CRA in conformity with the Memorandum of Understanding (MoU) as a Forum partner, and (3) contribute to the Forum Secretariat, either through in-kind contribution or through an annual fee (defined on a yearly basis to support Secretariat staff and activities). Typical members are ministries for science and research, scientific institutions, funding agencies, academies of sciences and research foundations. *Partners* are organizations that subscribe to the Belmont Challenge, but do not have a vocation to meet the criteria for membership.

Members and partners meet at least once every year in the Plenary Meeting, which is the highest decision-making level of the Forum. At present (July 2020) the Forum has 30 members and six partners. Prior to becoming a member, organizations can be invited by the SC to participate in the open sessions of the plenary as *Observer*.

The SC is elected by the Forum members to support and advise the Forum activities between plenary meetings. The committee currently consists of eight members headed by two *co-chairs*.

This overall governance structure of the Forum is overseeing and supporting specific governance bodies at the level of each CRA. *TPO* are established by one or more of the partner organizations' head offices that is participating in the CRA and is (supported by the Forum Secretariat) entrusted with the preparation, publishing and management activities during the period of a CRA implementation. *GPC* are responsible for the practical implementation of a given CRA. This group is composed of a mandated representative from each member/partner organization participating in the CRA and provides overall responsibility for his/her organization's involvement in the CRA. GPCs are open to any Forum member interested in the CRA but also to other funding organizations, which may join a CRA without becoming a Forum member.

The Forum Secretariat is the executive arm of the Forum and is tasked with carrying out decisions of Plenary Meetings and Steering Committee, including internal communication. The Secretariat also provides support to the CRAs in many ways: in the scoping phase, with ICT (BFgo), in the launch (e.g. answering questions submitted by proposers, putting out press releases), training (e.g. training of researchers in transdisciplinary research) and has a responsibility in monitoring the project progress and implementation (e.g. organizing progress workshops and valorization meetings). The Secretariat is currently hosted at the Inter-American Institute for Global Change Research (IAI), in Montevideo, Uruguay. Before 2018, the Secretariat was hosted at the French National Agency for Research (ANR), Paris, France.[9]

The Secretariat consists of a director and staff. The director is appointed by the SC for a period of three years. All other staff consists of in-kind contributions of personnel time (two full-time AAAS fellows contributed by the United States National Science Foundation (NSF) and 0.2 full-time equivalent (FTE) each from the European Commission (EC), the Natural Sciences and Engineering Research Council of Canada (NSERC), Thailand Science Research and Innovation (TSRI) and (more intermittently) the Ministry of Science, Technology and Productive Innovation in Argentina (MINCyT) and the (National Natural Science Foundation of China (NSFC)).

The roles of the various bodies as described in the ToR are summarized in Table 3.

Table 3 Roles and tasks for different permanent Belmont Forum governance bodies

Body	Role, election/appointment	Tasks and activities
Members	Contribute to the BF's mission and goals. New members need to be accepted by the Plenary Meeting	• Formally declare their commitment and willingness to actively contribute to Forum operations • Have participated in a CRA in conformity with the MoU • Contribute to the Secretariat (in cash or in kind)
Plenary Meeting	Main decision-making body of BF. Consists of principals from members who can take executive decisions regarding the mobilization of resources to support activities. Meets each year. Plenary Meeting needs 2/3 membership quorum	• Consider and adopt measures to establish, review, and update the policies and procedures of the BF, as well as to evaluate its work and the accomplishment of its objectives • Review periodically and approve the scientific policy and agenda of the Forum and consider and approve its long-range plans • Evaluate proposed CRAs and decide whether to be involved in scoping exercises • Pledge resources for approved CRAs • Evaluate Forum activities and recommend improvements • Elect the Members of the SC, including the co-chairs • Consider and approve the Rules of Procedure of the SC • Consider and approve the annual financial report and plan • Establish ad hoc committees or working groups as necessary
Steering Committee (SC)	Elected by the members. Supports, advises, and aligns the activities of the co-chairs. It is responsible for the implementation of decisions made at the annual meetings and makes decisions in between plenary meetings, unless the decision requires the full complement for adoption.	• Develop recommendations on the policies of the BF, for submission to and approval by Members • Ensure that the co-chairs and Secretariat implement the policies adopted by the Members and advise them accordingly, notably in cases where the co-chairs refer questions to the SC • Make recommendations to Members regarding Secretariat activities and the long-range plans and strategic planning • Make recommendations to Members regarding financial policies associated with supporting Secretariat functions and participating in CRAs

[9] Belmont Forum (2015). The Belmont Forum Terms of Reference

Body	Role, election/appointment	Tasks and activities
	Meets in-person at least twice a year in conjunction with the annual meeting of the members and virtually throughout the year as needed (but at least 4 times a year)	• Make recommendations to Members regarding amendments to the governance documents • Propose to Members the designation of new Members, Partners and Observers • Supervise the Secretariat; hire the director and staff • Perform any other functions entrusted to it by Members
Co-chairs	Elected by the members (for three-year terms renewable once) Each supported by an assistant from his/her staff, not part of the Secretariat	• To prepare the agenda of the Forum and SC meetings (supported by Secretariat) and to moderate and lead through the meetings' discussion from a neutral position with the objective to integrate different perspectives • To act as spokespersons of the Forum and to formally represent the Forum to external bodies, based on the positions formulated by the SC • To guarantee the continuous communication and information flows within the BF, assisted by the Secretariat
Secretariat	Primary executive element of the BF The Secretariat director is appointed by the SC for a 3-year period and reports to co-chairs and SC	• Implements the decision taken at annual meetings and by SC • Provides support to the CRAs in many ways: in the scoping phase, with ICT (BFgo), training (e.g. training of researchers in transdisciplinary research) and has a responsibility in monitoring the project progress and implementation (e. g. participating in weekly telecoms, helping with the launch, organizing progress workshops and valorization meetings) • Interacts through electronic communication means and meets physically at least once a year

Source: Belmont Forum Terms of Reference (2015)

3 Findings of the evaluation

3.1 Relevance

Key findings on relevance

- The Forum's vision, objectives and activities are still highly relevant today. The rationale for the establishment of the Forum is highly recognized at international political levels and the issue of global environmental change as well as the collaborative transdisciplinary approach of the Forum, which is at the core of the Forum action, have remained extremely important. While there are many forums and organizations that focus on environmental change, few of them offer active multilateral funding programs that specifically spur global transdisciplinary research.
- Vis-à-vis its members (which are mostly funding organizations), the Forum serves two relevant purposes – namely, (1) engaging a dialogue with the global research community on environmental challenges (discussion and potential alignment of research agenda), and (2) strengthening the transdisciplinary approach needed to address the complexity and global character of the problems. Both are key motivations to join the Forum.
- However, after 10 years in operation, the Forum is entering into a maturity stage. The increasing number and diversity of its members and the need for achieving a critical mass of members to support its operations, suggest that there is a need to put the strategic direction of the Forum on the agenda and to discuss more explicitly what the Forum wants to achieve in the future.

This chapter of the evaluation considers to what extent the Forum vision, objectives and activities still correspond to the underlying needs and problems related to global environmental change and achieving the SDGs. To examine the pre-existing problems and the way in which these have evolved over time and to look at the context in which the Forum operates and the conditions that are in place, we examined a range of background documents, strategic documents of the Forum, and policy documents in the GCR area. The information coming from these sources was supplemented by interviews with key stakeholders, surveys with members, partners, and beneficiaries, and the two in-depth case studies on the complementarity of the Forum with other initiatives in the field and the collaboration between the Global North and the Global South.

EQ1: How relevant are the rationale, mission, and activities of the Forum (in the context of understanding, mitigating, and adapting to global environmental change, as well as achieving the SDGs)? What developments can be expected in the next 10 years and how could the Forum anticipate on this?

To answer this first evaluation question, we distinguish between (1) the general relevance of the Forum in relation to global environmental change, (2) the relevance of the Forum in relation to its member and partner organizations, and (3) a discussion on how to ensure the appropriate objectives and strategic direction for the Forum going forward.

3.1.1 The continued need for international and transdisciplinary global change research

The Forum aims to respond to a highly complex and multi-layered challenge. As stated above, its main vision is to encourage *"International transdisciplinary research providing knowledge for understanding, mitigating, and adapting to global environmental change".* Moreover, the Forum pursues a number of specific goals that are linked to promoting and enhancing research

that supports the natural, social, and economic systems around global environmental change (incl. information provision, risk assessments, policy evaluations, studies and analyses).

The evidence from this evaluation suggests that **the Forum's vision, objectives, and activities are indeed still highly relevant today**. The relevance of research for adapting to and mitigating global environmental change was, among others, highlighted in the outcome document, *The Future We Want*, endorsed by all UN Member States at the UN Conference on Sustainable Development in 2012.[10] In 2017, the United Nations Environment Programme (UNEP) published a report on *Strengthening the Science-Policy Interface*, which stated:

> *"achieving the Sustainable Development Goals requires scientific advice on complex interactions between goals achievement, which are dynamic, non-linear and uncertain. Policy processes are also complex – with interactions of multiple parties producing uncertain outcomes."*[11]

This highlights a perceived need among the international environmental governance community for more science which transcends disciplines and involves stakeholders. The Global Sustainable Development Report 2019, published by the UN, under the auspices of the High-Level Political Forum (HLPF) for sustainable development, concentrated on the role of science in achieving the SDGs.[12] The report highlights that:

> *"the world now needs more sustainability science. That is a new, more engaged academic field of studies that sheds light on complex, often contentious and value-laden, nature-society interactions, while generating usable scientific knowledge for sustainable development."* (p. 120)

Sustainability science remains a niche field in research and will not become more common without *"significant adjustments to universities and other research and training institutions"*. **This report endorses the need for more inter- and transdisciplinary research on environmental change,** as well as the linkages with socio-economic processes**. The rationale for the establishment of the Forum thus still seems to be highly recognized at international political levels**, even as the political agenda has advanced from one with concepts and terms that are more distinct from a scientific perspective, such as global environmental change, to one that is more integrated, including the elaboration and endorsement of the SDGs.

In line with this international documentation, the evidence coming from the interviews with members and partners suggests that the issue of global environmental change and the collaborative transdisciplinary approach of the Forum has remained extremely important and has even become more pertinent over time. While there are many forums and organizations that focus on environmental change, few of them offer active multilateral funding programs that specifically spur *global transdisciplinary research*. Bringing the resources and expertise from different countries and different types of stakeholders at various levels together is considered crucial in generating the right knowledge and effectively informing policy.

Beneficiaries of the Forum also expressed positive views on the relevance of the Forum in relation to better understanding, mitigating, and adapting to global change. A large majority

---

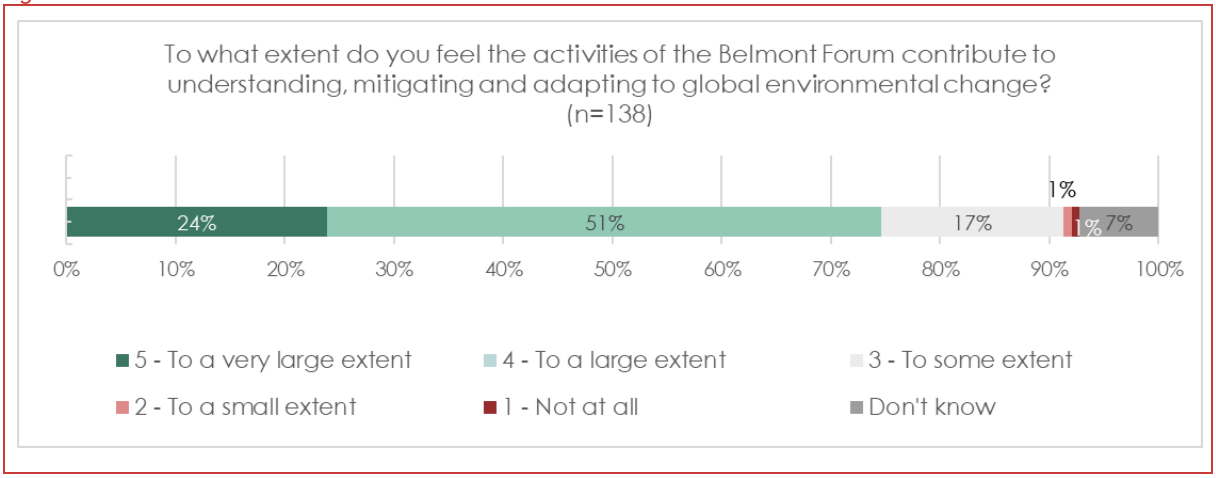
[10] Future We Want - Outcome document. Sustainable Development Knowledge Platform. Retrieved on 28.08.2020 from: https://sustainabledevelopment.un.org/futurewewant.html

[11] United Nations Environment Programme (2017). Strengthening the Science-policy Interface: A Gap Analysis. Nairobi.

[12] United Nations (2019). Global Sustainable Development Report 2019: The Future is Now – Science for Achieving Sustainable Development

of the surveyed beneficiaries **(75%) felt that the activities of the Forum were (very) relevant,** as shown in the graph below.

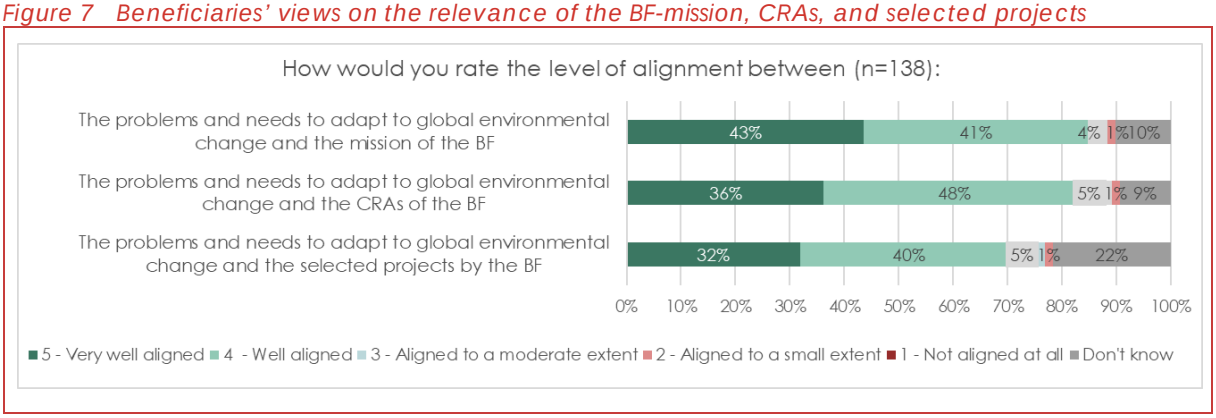
Figure 6 Beneficiaries' views on the relevance of funded activities



Source: Technopolis Group (2020)

When asked more specifically about the relevance of the Forum's mission, CRAs, and selected projects, beneficiaries also expressed very positive views. Between 72% and 84% of the surveyed beneficiaries were (very) positive of the alignment between the Forum's mission, CRAs, and projects and the problems and needs in the field.

Figure 7 Beneficiaries' views on the relevance of the BF-mission, CRAs, and selected projects



Source: Technopolis Group (2020)

3.1.2 Relevance of the Forum for funding organizations

Aside from the general relevance of the Forum, members and partners of the Forum were asked to elaborate on the relevance of the Forum's activities in relation to their respective organizations. Based on the interview and survey findings, it seems that most members felt that their membership to the Forum served a dual purpose.

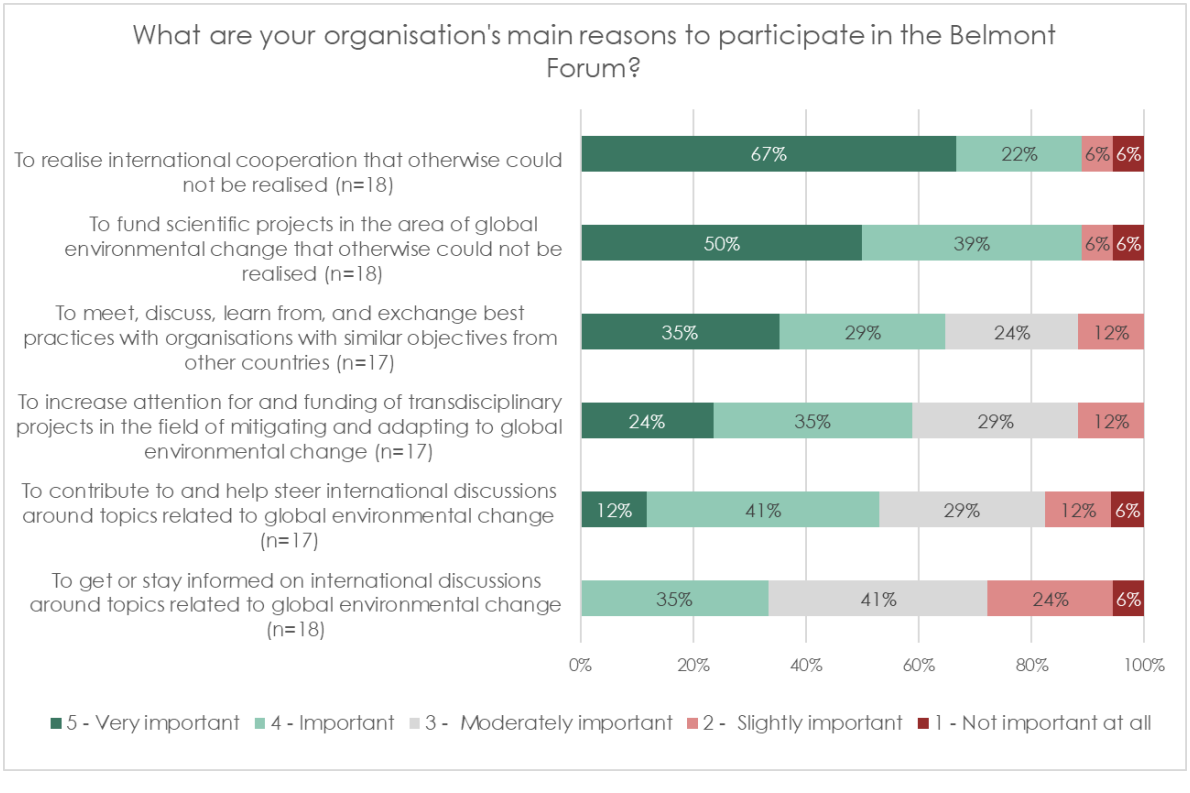
1. **To strengthen transdisciplinary research in global environmental change:** Many members and partners indicated that they primarily joined the Forum to support research on issues related to global environmental change. They argued that the overall objectives of the Forum or specific sub-areas [are?] very much aligned with the policy objectives in their own state, nation, or region. The transdisciplinary approach of the Forum was seen as highly

needed and appropriate to the complexity and global character of the problems to be addressed.

2. **To engage in international dialogue and action:** The alignment of funding agencies around the world was another frequently mentioned argument for the relevance of the Forum. A majority of members and partners mentioned that part of the reason for joining the Forum was to maintain international cooperation and relationships: They said they *"wanted to keep a toe in the water"*, *"joined because of the multilateral perspective and contacts"*, and *"wanted to have a link with the international community"*. Some even saw the Forum as an opportunity to make other funding agencies more aware of the urgency of research for environmental change or to change the way of thinking and working procedures of funders.

The online survey of members and partners shows similar findings. International cooperation and funding scientific projects in the area of global environmental change come out as the most important reasons for joining the Forum.

Figure 8 Members' and partners' main reasons for participation in the Forum



Source: Technopolis Group (2020)

3.1.3 *Strategic direction of the Forum going forward*

Having established the relevance of the Forum, both in relation to understanding, mitigating, and adapting to global environmental change and in relation to the objectives of members / partner organizations, it is important to consider what the Forum can and should do in order to ensure the appropriateness of its objectives and strategic direction going forward.

The evaluation findings suggest that the Forum has grown significantly in the past 10 years, and currently finds itself in a situation where important decisions must be taken in order to ensure a common understanding and commitment of members on what the Forum aims to achieve.

Members and partners of the Forum indicated that the Forum needs to find the right balance between the focus and ambitions of the Forum (e.g. what "transdisciplinary" research means, how many new CRAs should be started, and on the role of the Forum in valorizing research results), controlling Forum's agility (i.e. its ability to take decisions and move forward relatively quickly), which is at times somewhat hampered by the diversity in opinions between different members, and ensuring sufficient income to finance the central activities (e.g. by finding new members and keeping existing members on board).

One example of where members' opinions varied, concerns the Challenge (*Better understanding, mitigating, and adapting to global environmental change*). Several members indicated that this challenge is too broad. These members felt that a more focused mission might help the Forum to have a bigger impact. Some other members, however, felt that a broad definition of the Forum's main mission was necessary to keep the variety of members on board. Moreover, the latter group of members felt that there was no point in restricting the objectives and activities of the Forum unnecessarily. While both directions have their advantages and disadvantages, it is important for the Forum to discuss these explicitly and to come to a common understanding of the appropriate way forward.

3.2 Effectiveness and impact

Key findings on effectiveness and impact

- The Forum's scientific impact is impressive. In the bibliometric analysis, the Forum scored well ahead of several of the benchmarking groups and achieved similar results to the ERC (known for its focus on research excellence). The Forum made significant contributions to the scientific communities working on environmental change and associated topics, even above and beyond what might have been expected.
- The Forum also had a significant effect on policy, particularly scientific advice and evidence synthesis aimed at policymakers. The evaluation findings show that scientific publications funded by the Forum were cited quite often in science advice or policy documents. Almost a third of the publications funded in the period 2012–2014 were cited in one or more policy documents (e.g. from the EU, FAO, UNEP, and the World Bank). Data suggest that the Forum has provided meaningful input for science advice and argumentation by a range of national governmental agencies, IGOs and think tanks.
- In addition to the Forum's impact on the scientific communities and policymakers, it seems the Forum has also been relatively effective in disseminating knowledge and fostering discussion in wider circles and audiences. Forum publications scored well above the average world level for their OA accessibility (although lower than some benchmarking groups). Forum funding has contributed to an increase in OA accessibility in comparison to earlier publications by awardees. Forum-funded publications have been taken up and discussed by online and social media communities to a good degree, depending on the exact dimension considered. Forum publications performed best against the benchmark when it came to mentions in journalistic news items and on Facebook, did slightly less well on Twitter and were surpassed by a small group of benchmarking groups for Wikipedia citations.
- In terms of networking effects, the Forum has been effective in generating new collaborations and partnerships across actors, disciplines, and countries. Not only did the Forum have a very strong networking effect between funding organizations, it is also successful in fostering networking between researchers of different disciplines and countries, as well as between researchers and NGO-based partners.
- All in all, the Forum had considerable positive effects, which are likely to contribute to enhancing our understanding of global environmental change, as well as our ability to mitigate and adapt to this change.

The criteria "effectiveness" and "impact" of the Forum essentially come down to whether, and to what extent, the Forum's programs and initiatives have been realized and what effects and results have been achieved so far. The evaluation questions in this section look at the Forum's achievements from the multiple angles underlying transdisciplinarity – namely, its role in building networks and establishing new partnerships, the achievement of scientific outcomes, the dissemination of knowledge and outputs of the Forum to a wider audience, and any effects related to the design and implementation on policy. In this context, it is important to mention the myriad factors and stakeholders that affect our understanding of global environmental change as well as our ability to mitigate and adapt to this. This means that the evidence that we have been able to collect is not sufficient to simply attribute any progress in this area (or lack thereof) to the Forum only. Instead, the evidence gives us insight into the role that the Forum played within the context and complexity of issues related to environmental change as well as initiatives to better deal with it.

3.2.1 Engagement with key target audiences

EQ2: How successful has the Forum been in effectively engaging with its key (target) audiences?

This second evaluation question considers the extent to which the Belmont Forum interacts and engages with key target audiences effectively. The Forum's engagement with target audiences is a crucial condition to its overall effectiveness and impact, since it is at the basis of the Forum's visibility and reputation, and thus its ability to attract financial resources and high-quality proposals from strong consortia.

To engage effectively with its target audiences, the Forum developed a communication strategy (2016), with the primary goal to:

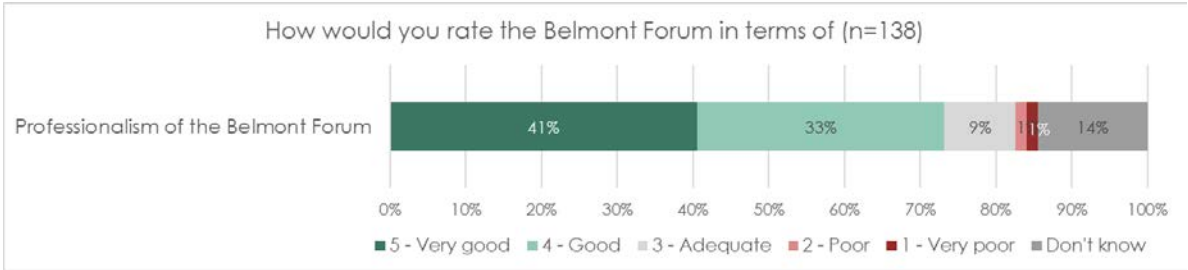
> *"strengthen engagement between research funding agencies, the academic research community and stakeholders and to improve coordination of early phase engagement on global change research strategies and priorities".*[13]

To do so, the communication strategy lays down a number of communication activities, namely to:

- "*Increase the awareness and knowledge among funding agencies and the scientific community about the Belmont Forum and its partnership approach, priorities and goals;*
- *Publicize CRA calls and conduct outreach to researchers, stakeholder communities and funding agencies to attract new organizations for possible membership in the Belmont Forum or partnership in CRA calls;*
- *Provide strategic and flexible communications support for value-added activities undertaken by the Secretariat, awardee meetings led by the TPOs, and events identified by members;*
- *Maintain transparency regarding Belmont Forum operations, membership and history while also ensuring the security of internal documents; and*
- *Support the work of Belmont Forum eI&DM* (e-Infrastructures and Data Management) *by closely coordinating our communication and engagement efforts with the C3O* (virtual Communication, Collaboration and Coordination Office)."

This chapter focuses on communication with applicants and beneficiaries. Engagement with its external stakeholders is discussed in the next chapter on networking, and engagement with members and partners is discussed in the chapter on efficiency and governance.

Figure 9 Beneficiaries' views on the professionalism of the Forum



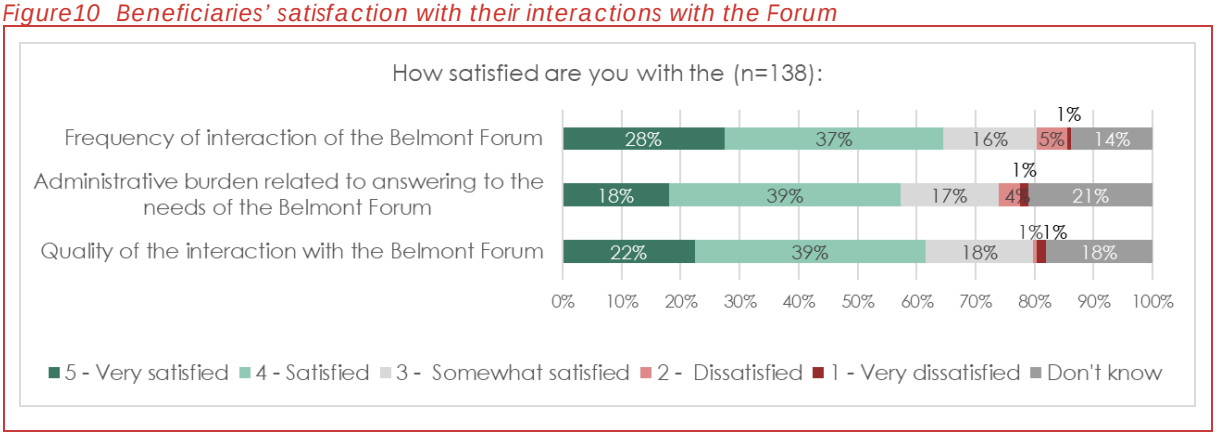
[13] Belmont Forum (2016). Belmont Forum Communication Strategy

Source: Technopolis Group (2020)

A total of 74% of beneficiaries felt that the level of professionalism of the Forum in general was either "very good" or "good". Only 2% provided a negative answer to this question.

However, when asked about the direct communications and interactions with the Forum, beneficiaries were somewhat less positive in their answers. Still, more than half of the surveyed beneficiaries were "very satisfied" or "satisfied" with the frequency and quality of their interactions with the Forum, and the administrative burden related to their funding of the Forum. Very few beneficiaries indicated to be dissatisfied.

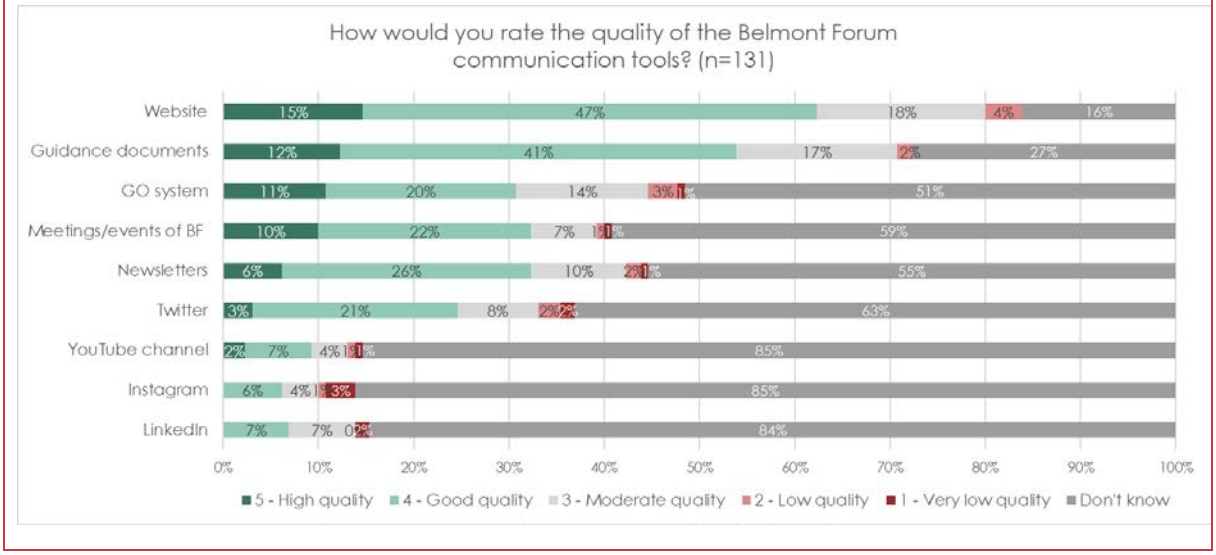
Figure10 Beneficiaries' satisfaction with their interactions with the Forum



Source: Technopolis Group (2020)

When asked about the communication tools used by the Forum to engage with their beneficiaries, respondents were most satisfied with the regular tools of the Forum, in particular its website and guidance documents. Beneficiaries were least familiar with the social media channels of the Forum (i.e. Twitter, YouTube, Instagram, and LinkedIn, as shown in Figure11).

Figure11 Beneficiaries' views on the quality of the Forum's communication tools



Source: Technopolis Group (2020)

As with the previous question, a relatively large number of beneficiaries were not able to answer these questions. In line with the evaluation team's own experience in engaging beneficiaries for the interviews and surveys, it seems that quite a few beneficiaries were not (actively) aware of the fact that they were a beneficiary of the Forum. Most likely, this was due to the fact that beneficiaries' primary interactions were with the TPOs or ERA-NET coordinators organizing the respective calls. Another cause for this may be the relative recent introduction of use of (some) social media (LinkedIn a couple of years ago, Twitter less than 2 years ago, Instagram only 3 month ago). Pre-2018 projects, especially, are not familiar with the social media channels at all and were only actively approached about them after May 2020 (so after the closing of the survey for this evaluation) when almost all of them had valorization activities.

The early stage of social media communication is also reflected in the results of our own desk research on the social media. As shown in the table below, the numbers of followers of the Forum accounts at various social media are quite low.

Table 4 Social media followers – retrieved on 1 July 2020

Communications channel	Number of followers
Facebook	86
Instagram (Belmont forum)	50
LinkedIn	372
Twitter (@Belmont_forum)	2477
YouTube	85

Source: Technopolis Group (2020)

To improve the communication activities, the Forum has been in the process of recruiting a dedicated communications director. This process has taken longer than foreseen, which may (at least in part) explain the somewhat limited online visibility of the Forum (especially to potential applicants and beneficiaries).

3.2.2 Networking effects

EQ3: Networking effects: How effective has the Forum been in generating new collaborations and partnerships across various sectors, disciplines, and countries around the globe? To what extent are the funded projects truly co-designed and co-created?

Whereas the previous question investigated the interactions between the Forum and stakeholders, this question looks at the extent to which the Forum incentivizes new interactions and collaborations between various types of stakeholders. In addition to examining the number and types of new collaborations, we also investigated to what extent projects were truly co-designed and, if needed, how the co-design process can be further encouraged in practice.

The answer to this evaluation question draws on a combination of the in-depth interviews and surveys of the Forum's members, partners, and beneficiaries, as well as the bibliometric analysis. The ensuing sections distinguish between (1) engagement with external partners and stakeholders, (2) the networking effects between funding agencies, and (3) the networking effects between researchers.

3.2.2.1 Engagement with external partners and stakeholders

In relation to the Forum's **engagement with external partners and stakeholders**,[14] it should be noted that we only spoke to a total of six partner organizations and five external stakeholders. These suggest that the Forum has been quite effective in engaging with external partners and stakeholders. Most interviewees indicated that the Forum has managed to build up valuable relationships with potential new members, partners of the Forum, and external stakeholders active in the field of GCR. The proactive marketing of the Forum, as well as personal contacts and relationships of members of the Forum, have contributed to this.

3.2.2.2 Improved networking between funding agencies

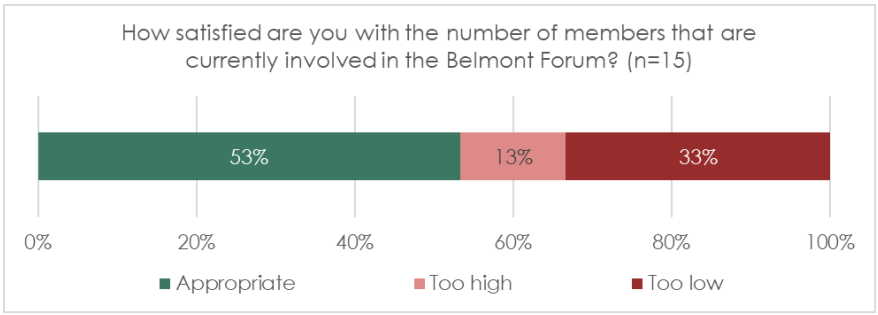
The findings from the interviews with members and partners show that the Forum has had a very strong networking effect between funding organizations. Members and partners indicated that the **Forum has played (and will continue to play) a very important role in the connections and collaborations between funding agencies**. In fact, members felt that the Forum helped them to better connect with their counterparts in other countries in various ways. For example, they indicated that the Forum helped them to:

- get to know their counterparts in other countries better (including the activities, working ways, personal contacts, etc.);
- bring a multilateral thinking into their organization;
- better alignment in activities and methodologies;
- learn from each other (for example in open data management processes);
- sell (new) ideas at home; and
- look beyond national interests.

These were all considered very valuable effects of the Forum.

When asked about the total number and the variety of members currently involved in the Forum, a small majority (53%) of the surveyed members felt that this was appropriate. A third (33%) felt that the number of members was too low, and thus needed to be increased.

Figure12 Members' views on the number of members currently involved in the Forum



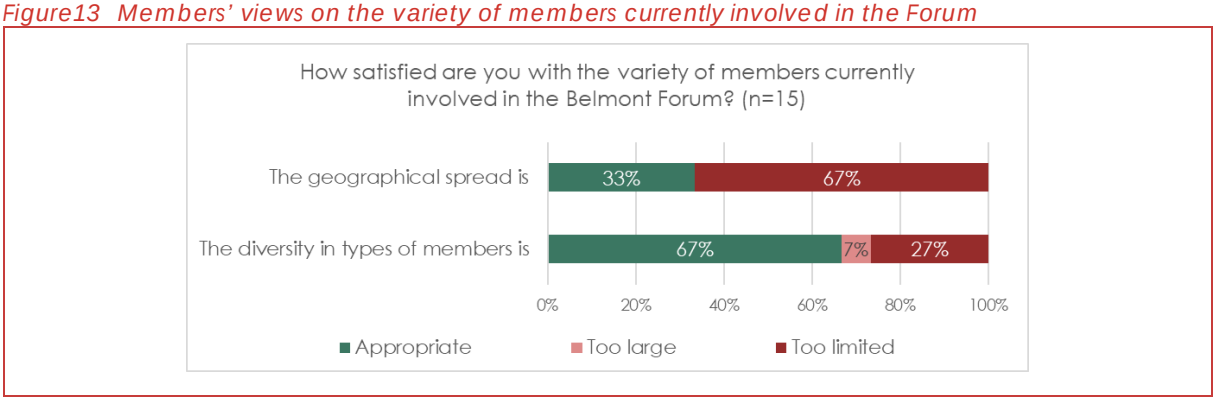
Source: Technopolis Group (2020)

The findings on the variety of members were less positive. While the general diversity was considered to be appropriate by most surveyed members (namely, 67%), two thirds of the

[14] These include other organizations active in the field of global environmental change (and who are not official members or partners of the Belmont Forum.

members felt that the **geographical spread of members was too limited**. This is in line with the interview findings, which reveal that a large portion of members, partners and beneficiaries feel that the Forum should include more members and beneficiaries from the Global South.

Figure13 Members' views on the variety of members currently involved in the Forum



Source: Technopolis Group (2020)b

3.2.2.3 Improved networking between researchers

In terms of networking effects between researchers, the Forum has been **effective in generating new collaborations and partnerships**. It has also been successful in fostering networking between researchers of different disciplines and countries and – even though the methodology for measuring it is very new and very labor intensive – networking between researchers and stakeholders from non-research backgrounds (an essential element of transdisciplinary research). The various indicators are described below.

International co-publication rates of the Forum, North–South collaboration

A first way to "measure" the international cooperation between researchers is to analyse the authors of publications (see Table 5). For this we analyzed the authors of the 2012–2014 CRAs of the Forum (research has already been progressing for some years for these CRAs, so a good number of publications have been published that have a long-enough citation window (3 years) for us to be able to examine the citations for these papers). **Almost 73% of these Forum publications were written as international co-publications** (their international co-publication rate (ICR) is 73%). Compared with other publications (supported by other programs or research councils) this is a high score. ERC-supported publications have an ICR of 80%. BiodivERsA and ANR papers have levels of 70%. NERC (64%), EC (63%) and BMBF (62%) score significantly lower. The Forum observation was well above world level in the thematic set (34%) and the combined main funders' measurement (44%).

The high level of international co-publication of Forum researchers can also be attributed to the participation of these researchers in Forum research, since their ICR before their participation in the Forum was much lower (50%), as is the rate in their present projects outside the Forum (59%). **Forum participation increases the international co-publications of participants.** Average numbers of countries and authors per **Forum publication were comparable to those of many funders** (2.7 countries and 7.3 authors).

The international co-publications of Forum papers **also show significantly higher North–South collaboration than other funders,** as shown by the portion of publications containing at least one author from an OECD ODA country and at least one from a non-ODA country (44%). NSFC was closest on this dimension with 33%. Neither BiodivERsA nor the ERC scored very highly on this dimension (10% and 29%, respectively). The hypothesis of a positive effect of Forum funding

on North–South collaboration is also supported by the lower shares observed for prior (27%) and parallel (36%) publications by Forum awardees (both leads statistically significant).[15]

[15] Although this finding may appear inconsistent with the above finding from the interviews that there is a need to bring more members and beneficiaries from the Global South, this is not the case. For instance, the positive bibliometric finding on North–South collaboration is not strictly limited to interactions between Forum beneficiaries as it also can also cover cases such as a Forum beneficiary from the North co-publishing with a non-beneficiary from the South. Plus, even if Forum promoted greater North–South collaboration, this does not mean it cannot do better, and the survey indicates a perception that it could do so.

Table 5 Belmont Forum's and benchmarking groups' achievements on networking effects, 2013–2019

Groups	Total N pubs	ICR	Avg N countries	Avg N authors	North-South ICR	II	$HIP_{10\%}$	MI	$HMP_{10\%}$
Belmont Forum, CRAs between 2012–2014	371	73% [68%|-|77%]	2.7 [2.5|-|2.9]	7.3 [6.5|-|8.1]	44% [39%|-|49%]	1.24 [1.20|-|1.28]	25.7% [21.3%|-|30.0%]	1.92 [1.76|-|2.08]	28.1% [23.8%|-|32.7%]
Matched benchmarking groups									
Non-BF publications by BF awardees	3,745	59% *	2.5 †	8.2	36% *	1.14 *	15.5% *	1.67 *	21.8% *
BF awardees prior publications	3,044	50% *	2.0 *	6.0 *	27% *	1.15 *	15.6% *	1.71 †	20.7% *
BiodivERsA, 2008 call	426	70%	2.7	7.1	10% *	1.11 *	14.1% *	1.49 *	20.6% *
Benchmarking groups from the thematic set									
World level	98,812	34% *	1.6 *	4.7 *	17% *	1.09 *	14.2% *	1.25 *	14.8% *
Selected funders combined	23,658	44% *	1.8 *	5.8 *	23% *	1.09 *	13.0% *	1.33 *	15.1% *
National Natural Science Foundation of China	6,392	35% *	1.5 *	5.6 *	33% *	1.11 *	14.4% *	1.32 *	13.1% *
National Science Foundation, US	4,062	45% *	1.9 *	6.1 *	18% *	1.08 *	13.1% *	1.34 *	16.0% *
European Commission	3,528	63% *	2.6	7.2	20% *	1.10 *	14.0% *	1.45 *	18.3% *
European Research Council	487	80%	3.5	11.1	29% *	1.14 *	16.3% *	1.45 *	14.0% *
Natural Environment Research Council, UK	1,431	64% *	2.6	8.2	21% *	1.09 *	15.3% *	1.50 *	19.7% *
Bundesministerium für Bildung und Forschung, DE	755	62% *	2.5	8.0	32% *	1.08 *	13.4% *	1.37 *	16.5% *
Agence Nationale de la Recherche, FR	595	70%	2.7	8.6	27% *	1.10 *	12.4% *	1.56 *	19.6% *

Note: Funders from the thematic set selected to include the 3 largest by output volumes (National Natural Science Foundation of China (NSFC); NSF; and EC), and the top 3 by $HIP_{10\%}$ scores. North–South ICR: international co-publication rate with a least one author from a country on the OECD ODA list and at least one author from a country not on the list. One-tail test for differences of means between the Forum and each benchmarking group: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

Interdisciplinarity and multidisciplinarity of publications

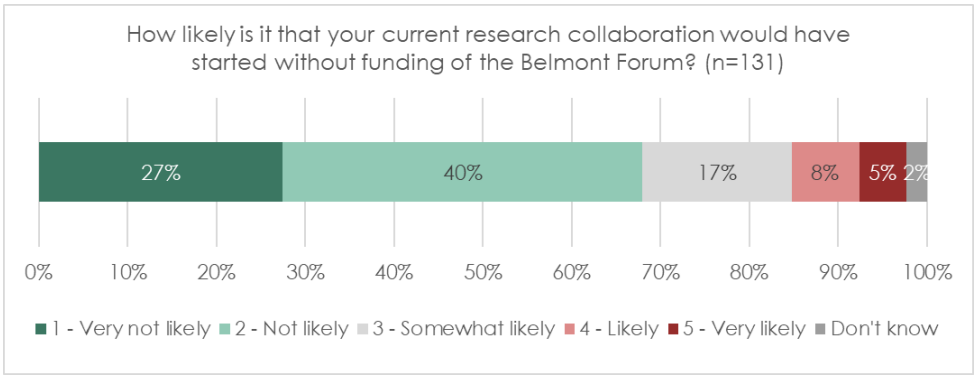
There is no methodology (yet) that can measure the amount of transdisciplinary research based on bibliometric data. However, transdisciplinary projects are often also interdisciplinary and multidisciplinary (among other dimensions). Indicators for Interdisciplinarity and multidisciplinarity are available. The interdisciplinary index is an indicator for interdisciplinarity, capturing the capacity to recombine knowledge from various disciplines within scientific papers. The Forum CRA 2012–2014 publications show an interdisciplinary index of 1.24, well above those of other funders. These findings show **(1) that Forum calls have been able to select highly interdisciplinary researchers, (2) that the Forum fostered a further increase of their level of interdisciplinarity, and (3) that some of the features of the joint calls** (also employed by BiodivERsA) **appear successful in fostering interdisciplinarity relative to less-specialized funding models** (i.e. relative to national funding agencies with a broader range of funding mechanisms).

Not only are many Forum publications interdisciplinary as such, **Forum publications that are interdisciplinary are also markedly more interdisciplinary than other interdisciplinary publications.** As many as 25.7% of Forum publications are within the top 10% of highly interdisciplinarity papers (where 10% would be the average). Again, Forum competitions selected highly interdisciplinary applicants but also improved the interdisciplinarity scores of the participants (as noted, these improved to 25.7%, whereas the interdisciplinarity score for non-Forum papers of the group of Forum participants remained at 15.6%).

The multidisciplinarity index (MI) measures the degree of diversity in the disciplinary background of a paper's co-authors. The disciplinary background of an author was assessed by the distribution of the author's publications across scientific subfields. Findings show once again that Forum CRA 2012–2014 publications reached the highest intensity on this dimension (1.92). Among the selected funders, ANR obtained the next best MI score (1.56), with BiodivERsA (1.49), and then EC- and ERC-supported articles (1.45 for both). Prior and non-Forum publications by Forum-supported investigators also returned higher scores than those of other funders (1.67 and 1.71). The implications thus remain that **Forum competitions have successfully selected highly multidisciplinary investigators and research collectives, and that the support offered itself contributed to further progression of Forum awardees on this dimension**. Looking at the 10% most multidisciplinary papers, very similar observations were made, with Forum achievements measured at a share of 28.1% of publications in this case.

These positive bibliometric findings were confirmed by the survey of beneficiaries conducted for this evaluation. As shown in Figure14, two thirds (67%) of the surveyed beneficiaries felt that it would not have been (very) likely that their (interdisciplinary and/or multidisciplinary) research collaboration would have started without funding of the Forum. Only 13% of surveyed beneficiaries ought this scenario either "likely" or "very likely".

Figure14 Beneficiaries' views on the likelihood of their research collaboration without Forum funding



Source: Technopolis Group (2020)

Publications with authors from different sectors

Besides interdisciplinarity and multidisciplinarity, another aspect of transdisciplinary research is cooperation across sectors. Since information on the sectors of authors in publications is not available in the large databases of publications such as Scopus, we manually coded authors of Forum awardees by sector of activity.

Among the CRA 2012–2014 papers, 96% mentioned at least one university-based author (see the "Descriptive statistics" section in Table 6). In total, 43% of publications included at least one author from a governmental agency. Based on the information we had, and the time available for this assessment, we could not definitively distinguish between authors at executive or legislative branches of government (policymakers) and authors from government-funded research centers. However, following a quick assessment, the vast majority of affiliations in this category appeared to belong to the second group. In all, 37% of Forum papers saw contributions from authors at large research centers, such as the Max-Planck institutes in Germany or the Russian Academy of Sciences institutes. A further 29% of Forum CRA papers had contributions from an author located at "other research centers" (ORCs), a category that included a dispersed set of independent research groups that could not be classified as belonging to the other groups. These centers may have been think tanks or research arms of philanthropic organizations, sometimes veering close to an NGO, but always with a clear scientific focus. **Authors from NGOs appeared on almost 9% of Forum-supported publications, by far the highest level of participation from NGO authors recorded in the comparison**. **Authors from intergovernmental organizations (IGOs) appeared on slightly more than 5%. Finally, authors from the private sector appeared on slightly less than 5% of papers.** It should be remembered in interpreting these results that some authors did hold affiliations crossing multiple sectors. It should also be remembered that it is in principle possible that additional or even most collaborations with NGO-, IGO- or private sector-based authors may have taken place in other project activities outside the preparation of peer-reviewed publications.

Forum publications were also examined to identify publications that combined multiple sectors, especially the three non-research sectors included in the analysis (see the "Research-society co-publications" section in Table 6). Forum **publications achieved higher percentages of inter-sectoral authorship (15%) than prior publications by Forum investigators (11%), but so did parallel non-Forum publications by these researchers.**

Forum collaborative publications appear to have included a fair number of NGO-based authors (almost 9%), significantly above parallel publications by Forum investigators (4.3%), prior publications (2.7%) and BiodivERsA (3.5%). Higher shares of Forum publications were also written

in collaboration with an IGO-based author (5.4%) than publications prior to funding (3.2%) or BiodivERsA publications (1.2%). Concurrent publications by Forum investigators had the same share of these publications involving an IGO-based co-author. Forum funding may not have been the deciding factor in fostering this specific type of co-publication for supported investigators, although the Forum competitions were successful in identifying investigators with good potential in this respect. Looking at co-publication with authors based in the private sector, Forum publications recorded a lower level on this dimension than prior publications by Forum investigators (4.3% to 5.9%). By contrast, the share of such co-publications increased in parallel papers by Forum investigators compared to the figure in prior articles (8.6%). BiodivERsA articles also showed a higher share of these co-publications (7.0%). It appears that **Forum support has contributed toward a shift in the focus of collaborative activity for supported investigators, from private sector partners to NGO-based collaborators**. This shift, however, was not accompanied by a net increase in collaborative activity since, as already shown above, the aggregate figure for collaborative activity irrespective of the exact sector (NGO, IGO or private) was roughly the same for Forum publications and concurrent papers by Forum investigators.

Table 6 Belmont Forum (and benchmarking groups) achievements in inter-sectoral collaboration, 2013–2018

Authorship	Belmont Forum, CRAs between 2012–2014	Non-BF publications by BF awardees	BF awardees prior publications	BiodivERsA, 2008 call
Descriptive statistics				
Research-side authors				
Academic research centers and	96.2%	94.3%	93.1%	94.8%
Governmental agencies	43.1%	48.4%	44.2%	51.9%
Large research organizations	37.1%	37.4%	32.5%	52.8%
Other research centers	29.0%	23.8%	17.9%	29.3%
Society-side authors				
Non–governmental organizations	8.7%	4.3%	2.8%	3.5%
Inter–governmental organizations	5.4%	5.5%	3.3%	1.2%
Private and business sector	4.6%	8.6%	5.9%	7.0%
Research–society co–publications				
Any research–society co–publication	15.4%	15.1%	10.6% *	11.0% †
Research & NGO	8.7%	4.3% *	2.7% *	3.5% *
Research & IGO	5.4%	5.4%	3.2% †	1.2% *
Research & Priv	4.3%	8.6%	5.9%	7.0%

Note: Shares of publications with a least one author with an affiliation in the category or categories of interest. Do note that authors may have multiple affiliations falling within multiple sectors (but a single author cannot alone amount to a co-publication). Large research centers: government-funded and quasi-academic centers, such as the Max-Planck network of institutes. Other research centers: research centers institutional status could not be clearly established but which appeared to be independent research centers; charity-based; or (less often) government-funded. IGO: intergovernmental organization. NGO: non-governmental organization. Priv: private. One-tail test for differences of means between the Forum and each benchmarking group: * p<0.01, † p<0.05, ‡ p<0.1. Significance testing was not performed on descriptive statistics.

Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

These bibliometric findings indicate that Forum projects have diversified co-publication patterns of participating researchers (to a certain extent). The exact implications of these findings for understanding team composition and research practices would require

conducting in-depth case studies of single projects, which was outside the scope of this evaluation.

3.2.3 *Scientific impact and excellence*

EQ4: Scientific outcomes: To what extent did the Forum contribute to the science base for environmental change (understanding, mitigation, and adaptation)?

This fourth evaluation question assesses the scientific outcomes of the Forum. More specifically, it focuses on what scientific outputs have been produced by projects funded by the Forum, and the impact of these outputs. Before answering this evaluation question, it is important to note that when consulting the Forum's various stakeholders about their views on the scientific impact, it appeared that many members, partners, and beneficiaries were unsure of the Forum's performance in this regard. Besides the fact that for many projects it was too early to assess the scientific impact, members indicated that (their view on) the monitoring and reporting of projects (at individual but also aggregated level) was too limited to develop a comprehensive overview and understanding of the Forum's scientific performance. For this, detailed narratives connecting the dots between stakeholders and describing processes are necessary. This was outside the scope of this evaluation, since the data collection for this is rather labor intensive (Schneider et al make the explicit recommendation that transdisciplinary programs employ a staff of dedicated "TD experts" to facilitate and track societal outcomes) and is only at an initial stage at the Forum.

As a result, the bibliometric analysis forms an important source of information for this question.

The projects funded by the Forum have contributed to the high scientific impact of the scientific communities working on environmental change and associated topics. This is demonstrated through Forum publications' high levels of citation impact, as well as through indicators such as the share of highly cited publications (as a reflection of research excellence), the prestigious journals in which Belmont-funded projects published their articles, and the high citation distribution index. Each of these elements are discussed in more detail below.

To start with, **publications supported by the Forum[16] were, on average, published within the most prestigious journals** (Weighted CiteScore (WCS) of 1.74), comparable with the ERC (1.76).

Publications supported by the Forum were also very well cited (compared to other publications in the same scientific subfields): well above those of benchmarking groups, and sometimes close to or above those of articles supported by ERC, a funder widely recognized for its focus on scientific excellence. Forum publications recorded an average of relative citations (ARC) of 2.70 (meaning 2.7 times the citations of an average publication in the same scientific subfield), only behind the ERC (3.19), and above the BMBF (2.37) and EC (2.41) publications (with non-significant leads).

[16] It should be noted that citation-based indicators have long been used as proxies of the impact (or influence) of scientific publications in assessing the value of their contributions to an expanding knowledge base. Nevertheless, it is important to consider that citation-based indicators as a whole rely on the assumption that citations are generally used to express intellectual debt, to point to the prior work on which one is building in generating research questions, observations, or methods, to take a few examples. However, citations are used for other purposes as well; in fact, citations are sometimes used to formulate critiques, which conveys the opposite of the positive ascription to citations that is tacitly assumed here. Citation-based indicators rely on the notion that critique-driven citations are much less frequent than impactful citations and therefore of negligible influence on bibliometric study results, so long as one is working with sufficiently large numbers. Of the 371 CRA 2012–2014 publications, 157 were published early enough for their citation window (i.e. the period over which they accumulated citations) to be long enough to allow for robust citation metrics. Citation-based indicators presented here (with the exception of the Weighted CiteScore, WCS) were based on this subset of 157 publications.

Table 7 Scientific impact of Forum contributions in expanding the science base for environmental change, 2007–2019

Groups	Total N pubs	N pubs citable	ARC	$HCP_{10\%}$	$HCP_{1\%}$	CDI	CDC	WCS
Belmont Forum, CRAs between 2012–2014	371	157	2.70 [2.2|-|3.3]	40.1% [32.7%|-|47.7%]	6.4% [2.5%|-|10.2%]	26.4 [22.2|-|30.6]		1.74 [1.60|-|1.90]
Matched benchmarking groups								
Non-BF publications by BF awardees	3,745	2,195	2.09 †	24.5% *	4.4%	19.6 †		1.45 *
BF awardees prior publications	3,044	2,755	2.26 ‡	27.3% *	4.2%	19.8 †		1.43 *
BiodivERsA, 2008 call	426	405	2.26 ‡	29.9% †	4.2%	27.5		1.61 ‡
Benchmarking groups from the thematic set								
World level	98,812	48,782	1.33 *	15.2% *	1.7% †	8.4 *		1.15 *
Selected funders combined	23,658	10,038	1.81 *	23.4% *	3.0% †	20.1 †		1.39 *
National Natural Science Foundation of China	6,392	2,511	1.50 *	18.7% *	1.8% †	14.7 *		1.20 *
National Science Foundation, US	4,062	1,835	2.09 †	28.2% *	4.7%	23.6		1.56 †
European Commission	3,528	1,463	2.41	31.3% †	5.4%	27.8		1.54 *
European Research Council	487	185	3.19	36.5%	10.7%	33.3		1.76
Natural Environment Research Council, UK	1,431	774	2.28 ‡	29.1% *	5.0%	23.9		1.64
Bundesministerium für Bildung und Forschung, DE	755	296	2.37	32.7% ‡	3.8%	27.9		1.48 *
Agence Nationale de la Recherche, FR	595	262	2.04 †	25.4% *	3.1% ‡	25.2		1.56 †

Note: N pubs citable: Number of articles in the benchmarking group's publication set for which the minimal citation window for computing robust citation indicators has been reached and for which citation-based indicators can be computed. One-tail test for differences of means between the Forum and each benchmarking group: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

When articles belong to the group of 1% most cited publications in their subfields, it is generally seen as a reflection of their large contribution to scientific excellence and/or their breakthrough character. **No less than 6.4% of Forum publications fell within this exceptional group of highly cited papers** (where 1% would be the average). This is a lower score than ERC papers (10.7%) but above those of the EC (5.4%), NERC (5.0%), NSF (4.7%), BiodivERsA (4.2%), and BMBF (3.8%) publication sets (not achieving statistical significance, however). Based on the 10% most cited papers, publications funded through the Forum came out first, with a proportion of 40.1%, a larger share than in the ERC publication set (36.5%).

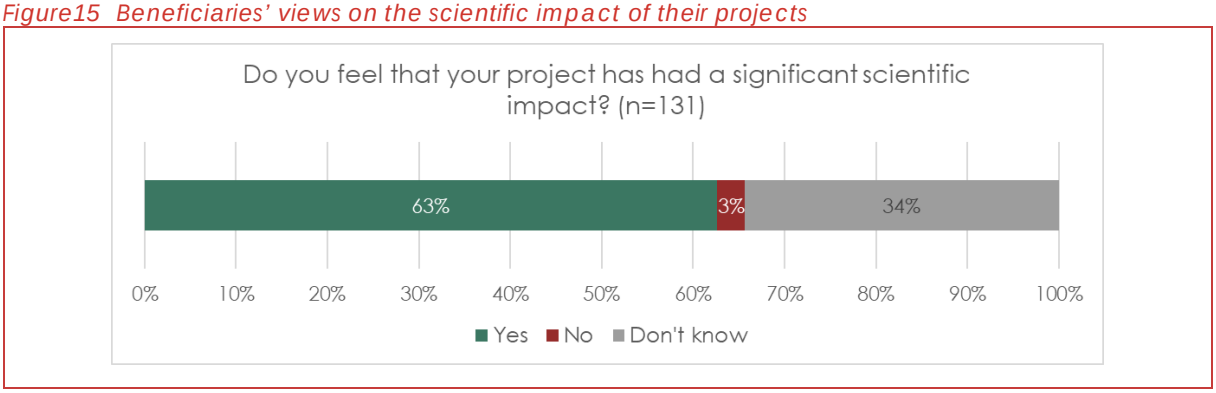
The citation distribution index (CDI) is another measure of impact accounting for all publications, one which is less sensitive than the ARC to highly cited publications. The findings of the analysis indicate that **the CDI of Forum publications (26.4) was much above world average** and similar to other funders in the set of selected benchmarking groups. The difference in the placement of the Forum based on the ARC and the CDI is attributable to the strong scores of the Forum's highly cited publications, which pull the ARC up. Given the small number of CRA 2012–2014 papers that could be used in computing these metrics, it is our view that the CDI provides a better reflection of "average" performance here.

The high CDI of the Forum – and of ERC, BiodivERsA, BMBF and EC – shows that, overall, it has a very strong research influence. For example, there is a strong concentration of CRA 2012–2014 papers in the highest citation deciles, with a majority of them (64%) falling in the top three deciles in which the share is always above expectations (as revealed by the green bars to the right of the citation distribution chart (CDC, a visual depiction of the CDI). As a consequence, the number of papers in the three least cited deciles is much less (6%) than expected (30%; revealed by the long red bars to the left of the CDC).

The Forum successfully selected highly influential scholars and likely enabled further improvements to their scientific impact: papers produced by Forum investigators prior to their Forum awards and concurrently with their Forum awards (i.e. non-Forum papers) often reached significantly lower citation impact levels than Forum publications (with exceptions for the $HCP_{1\%}$; and the ARC for prior publications). While prior and non-Forum publications by Forum investigators displayed measurements below those of the majority of the selected funders, their scores remain well above world level. This points to the Forum being successful in selecting outstanding scientists through its CRA calls, in addition to making a real contribution toward building capacity for high citation impact research.

The positive findings on scientific impact were confirmed by the – albeit limited – qualitative evidence that we were able to retrieve through the evaluation. Members and partners of the Forum felt that the projects that they were familiar with produced high-quality outputs and publications. Beneficiaries of the Forum also expressed positive views on the scientific impact of their projects. Two thirds of the respondents (63%) felt that their project had a significant scientific impact, whereas 34% did not know (yet).

Figure15 Beneficiaries' views on the scientific impact of their projects



Source: Technopolis Group (2020)

3.2.4 *Dissemination of research results and open data policy*

EQ5: Wider dissemination of knowledge: How effective has the Forum been in disseminating knowledge and other outputs generated by the Forum? To what extent were results of the Forum disseminated, taken up and discussed beyond academic circles?

This question considers the wider dissemination of knowledge generated through projects funded by the Forum. This section of the evaluation assesses the accessibility and dissemination of both scientific and non-scientific outputs and information to various types of key stakeholders outside the academic world. To answer this question, the evaluation draws on findings from the desk study, bibliometric and altmetric analysis, the interviews and the survey of beneficiaries.

Additionally, the Forum has invested significant time, effort and expense into developing an advanced OA policy, and related tools and activities, to promote the dissemination of both Forum publications and non-journal project outputs. The ramifications of this policy on project-level outputs could not be systematically assessed, but a few of the findings obtained provide a preliminary indication as to its impacts.

To summarize already the findings that will be presented below, **Forum publications scored well above the average world level for their OA accessibility, and notably above prior and non-Forum publications by awardees (although lower than some benchmarking groups).**

Forum-funded publications have been taken up and discussed by online and social media communities to a good degree, depending on the exact dimension considered. Forum publications performed best in the benchmarking exercise when it came to mentions in journalistic news items and on Facebook, did slightly less well on Twitter, and were surpassed by a small group of benchmarking groups for Wikipedia citations. Turning to web citations such as blog posts, videos, policy briefs or research tools and data sets, the main conclusion is that impact appears to have been achieved but that definitive evidence of these achievements could not yet be collected.

3.2.4.1 Online attention toward journal-based outputs

Care by researchers to make their findings available within OA peer-reviewed publications is increasingly used an indicator of participation in open science practices. OA availability can be considered as a facilitating factor in fostering online attention toward peer-reviewed publications. CRA 2012–2014 publications were published under an OA modality in a proportion of 63.6% (see Table 8). This observation was below levels recorded for publications supported by NERC (82.2%), BiodivERsA (73.6%), and ERC (71.8%). However, the Forum's score was above or well above the remaining funder publication sets, and well above world level and the main

funders' combined figure. **The Forum support appeared to have increased awardees' propensity to publish with an OA license** (scores for papers published by Forum awardees prior to and concurrently with their Forum-supported papers being seven and six percentage points below the Forum figure).

Mentions of a peer-reviewed publication by a

Forum-supported project on (social) media are an indicator for the spread of their content outside the scientific world. For this study various mentions of Forum publications and of publications from the various benchmarking groups in journalistic news items (with a known bias toward English-language and Northern sources), on Wikipedia, on Twitter and on Facebook were obtained from the PlumX database (a database of "the ways people interact with individual pieces of research output (articles, conference proceedings, book chapters, and many more) in the online environment") and compared on various altmetric indicators.[17]

Forum funding has had a significant positive effect on mentions of the peer-reviewed publications of the researchers of the CRA 2012–2014 in journalistic news items: the **Forum CRA 2012–2014 publications are mentioned more than three and half times (3.60) more often in journalistic news items than the average scientific publication,** and also more often than publications within the panel of benchmarking groups. In all, 24% of Forum publications belong to the 10% most mentioned papers in the journalistic news items, and 6.8% to the 1% most mentioned.

Wikipedia citations, a proxy for educational impact, for Forum-supported publications are around average level (e.g. 12.4% on the 10% most cited publications; here, BiodivERsA scores best, with15.7%).

The share of Forum publications receiving at least one mention on Twitter was 60% above world level (AMI = 1.61), a level below that of most benchmarking groups (BiodivERsA (2.63), NERC (2.08), ERC (1.89) and NSF (1.68) publications). However, Forum publications registered the second highest observation (4.9%), just below the ERC (5.2%) on the 1% most mentioned publications. This strong score for the Forum suggests that the instances when Forum papers stood out on this dimension cannot be purely attributed to self-promotion, as can often be the case with Tweets. For all indicators, Forum funding appeared to lead to specific increases in awardee achievements on this dimension.

Forum publications received high levels of attention on Facebook, about twice the world level. This is at comparable levels to the benchmarking groups, as was the case for the share of the 10% most mentioned articles (21.3%) and share of 1% most mentioned articles (4.0%). Again, these scores amounted to increases because of awardees' Forum participation.

[17] We invite the reader to closely read the corresponding section on methods and limitations for these strategies in the Appendix F.

Table 8 Belmont Forum (and benchmarking groups) achievements in disseminating knowledge and attracting online attention, 2007–2018

Groups	Total N pubs	OA%	News			Wikipedia			Twitter			Facebook		
			AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$
Belmont Forum, CRAs between 2012–2014	371	63.6% [59.0%|-|68.9%]	3.60 [2.89|-|4.31]	24.2% [20.3%|-|28.1%]	6.8% [4.3%|-|9.4%]	2.56 [1.39|-|3.90]	12.4% [10.7%|-|14.4%]	2.1% [1.1%|-|3.3%]	1.61 [1.48|-|1.73]	25.1% [21.0%|-|29.5%]	4.9% [2.7%|-|7.3%]	1.98 [1.63|-|2.33]	21.3% [17.4%|-|25.4%]	4.0% [2.2%|-|6.2%]
Matched benchmarking groups														
Non-BF publications by BF awardees	3,745	57.1% *	1.87 *	13.9% *	2.4% *	1.63 ‡	11.0% ‡	1.6%	1.39 *	19.6% *	3.0% ‡	1.50 *	14.8% *	1.8% †
BF awardees prior publications	3,044	58.1% †	1.43 *	10.7% *	1.6% *	1.25 †	10.8% ‡	1.2% ‡	1.25 *	12.0% *	1.2% *	1.29 *	11.0% *	1.2% *
BiodivERsA, 2008 call	426	73.6%	2.09 *	12.6% *	2.8% *	3.30	15.7%	4.0%	2.63	27.3%	3.2%	1.74	15.1% *	1.4% †
Benchmarking groups from the thematic set														
World level	98,812	43.9% *	0.91 *	9.5% *	0.9% *	0.95 †	9.9% *	0.9% †	0.93 *	9.7% *	0.9% *	1.01 *	10.1% *	1.0% *
Selected funders combined	23,658	47.6% *	1.41 *	12.1% *	1.6% *	1.29 †	10.5% †	1.3% ‡	1.23 *	14.3% *	1.4% *	1.27 *	12.5% *	1.3% *
National Natural Science Foundation of China	6,392	37.4% *	0.41 *	6.9% *	0.3% *	0.28 *	8.9% *	0.3% *	0.62 *	2.3% *	0.3% *	0.50 *	4.1% *	0.4% *
National Science Foundation, US	4,062	55.8% *	2.82 †	19.5% †	3.4% †	2.29	12.0%	2.5%	1.68	24.2%	2.9% ‡	1.59 †	16.1% *	1.7% †
European Commission	3,528	54.2% *	1.54 *	12.8% *	2.0% *	1.54 ‡	10.9% ‡	1.6%	1.44 *	19.0% *	2.2% †	1.77	17.9% ‡	1.2% *
European Research Council	487	71.8%	2.94 ‡	20.4% ‡	5.2%	3.21	13.2%	3.3%	1.89	32.7%	5.2%	2.08	21.0%	3.0%
Natural Environment Research Council, UK	1,431	82.2%	3.26	20.8% ‡	4.1% †	2.76	12.8%	2.7%	2.08	40.6%	4.5%	1.95	19.2%	1.5% †
Bundesministerium für Bildung und Forschung, DE	755	43.0% *	1.57 *	12.9% *	2.5% *	1.33 ‡	10.3% †	1.4%	1.25 *	14.4% *	1.6% *	1.61 †	15.1% *	1.1% *
Agence Nationale de la Recherche, FR	595	52.4% *	1.11 *	10.6% *	1.5% *	1.76	11.4%	2.0%	1.52	19.4% †	2.0% †	1.56 †	15.3% *	1.7% †

Note: OA%: Share of papers made publicly available under an open access publication mechanism. One-tail test for differences of means between the Forum and each benchmarking group: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus and PlumX databases (Elsevier)

In summary, it can be concluded that **scientific publications funded by the Forum have to a good degree been taken up and discussed by online and social media communities** and have reached audiences beyond academic circles. This is particularly evident from the high number of mentions in news items (the highest here, in fact) and Facebook mentions. High measurements on shares of publication in OA, as well as Wikipedia and Twitter mentions, were also recorded, although here Forum publications were surpassed by three or four benchmarking groups. Using a counterfactual, Forum funding also appears to have promoted an increased uptake of the outputs of its awardees in the news, Twitter and Facebook, and to have increased OA availability, thereby likely contributing to an increased societal impact of their research outputs. A similar effect was very likely for Wikipedia mentions as well, although the statistical robustness of findings was not definitive here.

3.2.4.2 Online attention toward non-journal outputs

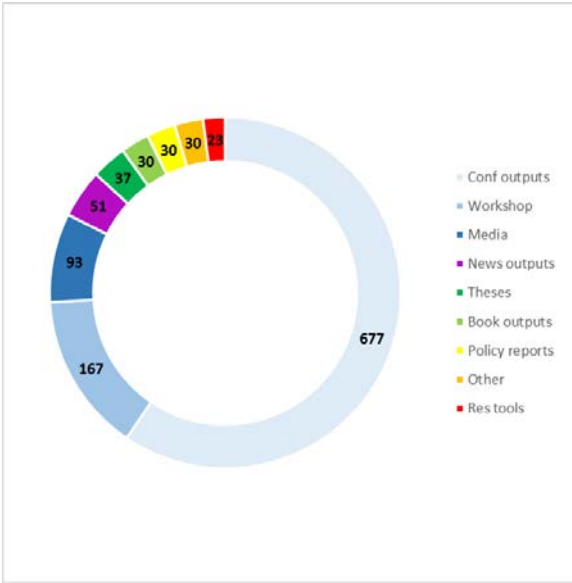
Types of non-journal outputs fostered by the Forum

Given the general challenges currently facing quantitative assessments of societal outcomes of research, as well their intermediaries, we have opted to rely on a mixed-methods approach combining multiple sources (BFgo, manual querying of project websites, as well as commercial portals recording hyperlinks toward webpages) to obtain a better portrait of achievements of the Forum on these dimensions.

Restricting the analysis to project outputs that could not be matched to the Scopus database, a total of 1,138 additional output entries were retrieved. These outputs were produced by projects funded through the CRAs 2012 to 2016.[18] They fell into nine broad categories, presented in Figure 16.

[18] Although note that BFgo does not contain records for CRA 2012: entries here were retrieved from project websites only. The analysis was restricted to the outputs produced by projects funded through the CRAs 2012 to 2016 because of the general research evaluation good practice to allow funded projects a fair amount of time for their outputs and outcomes to be realised (and as a consequence the use of this subset for the bibliometric analysis as well), and the extensive manual curation needed on answers from awardee reports before they could be used, given sometimes incomplete information that needed to be triangulated with online queries.

Figure 16 Belmont Forum achievements in fostering non-journal or non-academic outputs, 2012–2020 (projects from the 2012–2016 CRAs)



Note: Conf outputs include conference presentations, poster presentations, session organization and other academic conference-related activities. "Workshops" included only non-academic events with some degree of stakeholder or public engagement. "Media" includes social media and video (YouTube, Vimeo) outputs. The "other" category included working papers, confidential reports, newsletter issues. "Res tools" includes publicly available data sets, modeling tools and other research tools.

Source: Prepared by Science-Metrix using the BFgo database and web queries.

Web-citation impact of non-journal outputs

In this section, the extent to which these non-journal outputs were cited online is discussed (which is of course only one indicator for the uptake of research results outside the scientific world). First, the findings show that **a moderate share of Forum project websites received a web citation (see Table 9).** In total, 30 out of 49 CRA 2012–2016 websites had been the target of at least one hyperlink (61%). A greater share of BiodivERsA project websites had received at least one web citation (83%). Looking at the mean and median web-citation levels of those websites that had received at least one hyperlink, Forum and BiodivERsA project websites appeared to perform at roughly similar levels. BiodivERsA did somewhat better on the mean (13.4 versus 10.2 for the Forum) but slightly worse on the median (8.5 versus 9.0 for Forum websites). The low share of Forum project websites to have received a web citation points toward a target for future improvement (websites are often unfunded and may disappear after a project is closed, but the Secretariat now offers for projects to transfer any resources—documents, PowerPoints, etc.—from their closing project website to the Forum website's Resources page, which is considered a good initiative by the evaluators).

In other research outputs made available online, Forum outputs fared better. Looking at shares of outputs to have received at least one web citation, Forum outputs were ahead of BiodivERsA on this dimension in all four main categories of work considered. **The Forum's lead was clear for social media and video outputs**, with a share of 37% of such works cited at least once online, against 5% for BiodivERsA. For journalistic news outputs, Forum content was hyperlinked in 56% of cases, compared to 25% for BiodivERsA. Shares of content hyperlinked once or more was 67% (Forum) versus 50% (BiodivERsA) for research tools, and 41% (Forum) versus 36% (BiodivERsA) for policy reports.

Turning to mean and median levels of web citation within the subset of outputs that has received at least one hyperlink, BiodivERsA was clearly ahead for web-citation levels toward research tools, with a mean of 31.5 (against 3.5 for Forum) and a median of 14.5 (against 3.0 for Forum). The Forum was ahead for its journalistic news outputs, with a mean of 9.3 (against 3.0 for BiodivERsA) and a median of 5.0 (against 3.0 for BiodivERsA). The Forum recorded higher observations for mean level of citations in the categories of social media (2.8 to 2.0) and policy report outputs (20.7 versus 13.0), but was behind on the median in the same categories (1.5 to 2.0 on social media; 2.0 to 3.5 on policy reports). This discrepancy indicated that hyperlinks to Forum outputs in these two categories were skewed toward a few items, whereas they were slightly more evenly distributed in the case of BiodivERsA outputs.

Overall, **the picture that emerges from this web-citation analysis is that Forum-supported dissemination outputs were effective in finding an online audience, whereas project websites were not so**.

Table 9 Forum achievements in fostering online attention toward non-journal outputs, 2012–2020

Type	Funder	Online count	Linked 1+	Share 1+	1+ mean	1+ median
Project websites	BF	49	30	61%	10.2	9.0
	BD	12	10	83%	13.4	8.5
Media	BF	91	34	37%	2.8	1.5
	BD	19	1	5%	2.0	2.0
News outputs	BF	43	24	56%	9.3	5.0
	BD	4	1	25%	3.0	3.0
Policy reports	BF	17	7	41%	20.7	2.0
	BD	11	4	36%	13.0	3.5
Research tools	BF	12	8	67%	3.5	3.0
	BD	8	4	50%	31.5	14.5

Notes: Forum: Belmont Forum. BD: BiodivERsA. Online count is the count of outputs in the category that were made available online. Because not all outputs have been made available online, this count is often lower than the total count previously reported for that output category. Share 1+: share of output counts hyperlinked at least once. 1+ mean and 1+ median: mean and median of hyperlinks received within the subset of outputs with at least one hyperlink. Hyperlink counts are subject to multiple limitations, please see the methods section.

Source: Prepared by Science-Metrix using the BFgo database, Uber suggest and web queries

The box below presents the findings of several small case studies that were conducted on the web impact, stakeholders, and outcomes achieved by different Forum projects.

Case studies of web impact, stakeholder, and outcomes achievements

In terms of web impact, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) 2018 *Assessment Report on Land Degradation and Restoration*, to which DEVIL member Robert Scholes contributed as an editor and author, received a considerable amount of online attention, the most seen as part of this study. However, the writing of this report was a highly collaborative endeavor, and so we have not considered it to be an adequate example of the specific contributions made by Forum-funded initiatives. Next behind the IPBES report, Pan-Arctic Options' investigator Paul Berkman's article published in the *Science & Diplomacy* trade journal has also seen considerable attention, including from the Wikipedia page on the concept of the "Arctic". Most of the online attention recorded can be traced back to the original Wikipedia entry, which has been reproduced in entries on wiki websites that derive their content from Wikipedia. An opinion piece published by Berkman in the *Alaska Dispatch News* was also moderately hyperlinked (5 hyperlinks).

The Deltas project produced a YouTube video, "Why Do Rivers Have Deltas?", that received 25 hyperlinks, including from the Hebrew-language entry for "deltas" on Wikipedia. A *Nature Climate Change* publication by the project team also attracted attention in journalistic outlets after an initial write-up by *Scientific American*. The *Scientific American* piece was hyperlinked 17 times, although online attention was fueled in part by a controversy surrounding the tone of the coverage employed by a *New York Times* editorialist.

Members of the Urbanising in Place project (Nexus 2016 call) have contributed to the launch of the Citizen Soil Clinic network in London. This initiative will serve as a platform for soil-oriented citizen science database. These team members have also engaged municipal stakeholders (Greater London Authority) as well as conservation-oriented charitable organizations. Project members in Argentina and Belgium also interact with local municipal actors, in Rosario and Brussels respectively. Project members have been active in diffusing the outcomes from their interactions and workshops through journalistic interviews. Their proposal for the formation of a new Centre for Agroecology in Brussels has seen some online attention (8 hyperlinks).

The In-Source project has contributed to energy transition and water management practices in New York City, Vienna and different locations in Germany. In New York, team members have supported the municipal Department of Environmental Protection and the local chapter of the American Institute of Architects in planning for rising waters and for wastewater management. A New York Institute of Technology news release on the project attracted the attention (e.g. hyperlinks) of publications such as a *Scientific American* blog, which was in turn hyperlinked 20 times, by sources including *American Infrastructure* magazine.

The NILE-Nexus project (Mountains 2015) has modeled the Blue Nile's river flow, providing findings of direct relevance to at least four Ethiopian public sector partners, including the Ministry of Waters Resources, Irrigation and Energy. A blog post with contribution from team members was hyperlinked six times, including by Ethiopian news outlets.

3.2.4.3 The Forum's open science policies

The Forum recognizes the crucial role of open and effective data[19] and information exchange to the Belmont Challenge and therefore works to coordinate and promote access to transdisciplinary research data. To promote the use of research results, and therefore directly contribute to the aim of international research cooperation, the Forum has developed an *Open Data Access* policy, that was already adopted in 2015 (before 2016, when the FAIR data principles were published, to which projects funded by the Forum must adhere). Across all CRAs, the Forum e-Infrastructure and Data Management Team and Secretariat, TPOs and GPCs, work to implement *open data policy and principles*.

The high shares of Forum-supported publications with OA availability, and especially the increase measured relative to prior and non-Forum projects, indicate clear impact on this dimension. It is, however, impossible to establish a direct causal link to the Forum's OA policy, especially given how it was launched in 2015, and many of the projects whose publications have been evaluated were awarded their funding between 2012 and 2014.

The analysis of non-journal outputs of supported projects already presented above captured a certain number of research tools and data sets made available publicly online. Of particular interest here, and considering projects financed by the 2012–2016 CRAs, 7 out of 75 projects

[19] This also covers digital objects, such as software and code.

were found to have made data sets available online (mentioned in awardee project reports or as identified in project websites, and for which the availability of the data set online could be validated).

3.2.5 Policy impact

EQ6: Policy effects and outcomes: To what extent did results of the Forum foster policy debate or developments at international and national level or facilitate policymaking / implementation?

Evaluation question 6 zooms in on policy debate and policy effects. This assessment benefited from survey answers and the analysis of awardee report answers. The greater part of this assessment is informed, however, by the use of quantitative citation analysis drawing on a novel database, Overton.

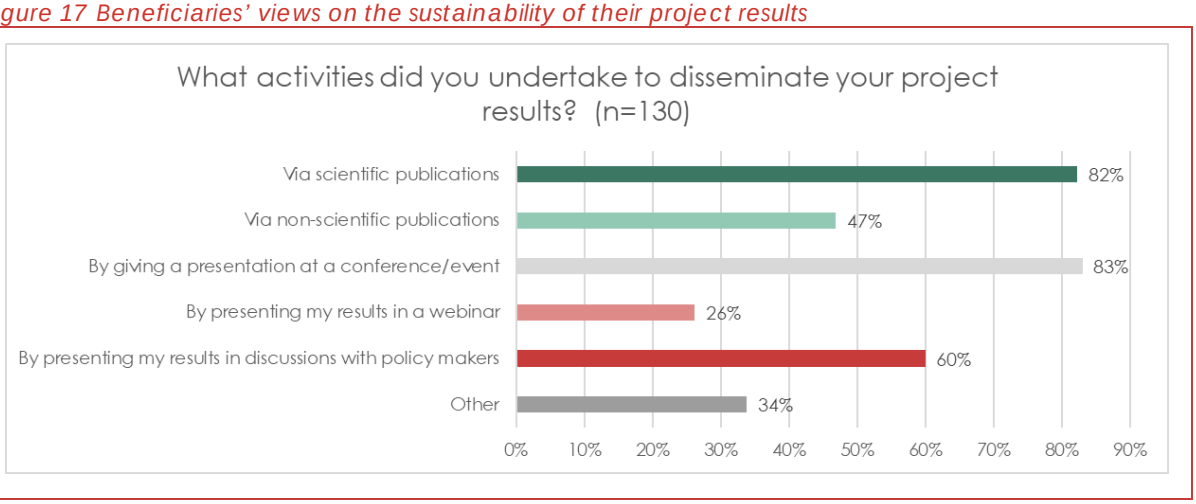
Of the respondents to the beneficiary survey, 60% indicated that they presented their results in discussions with policymakers, which is a confirmation of the self-reported project outcomes.

For those publications from CRA 2012–2014 projects, close to a third were cited by at least one science advice or policy document. This indicates that the research had been useful input, at least into the first steps in the process of knowledge transfer toward policy and regulation. Notable policy citations originated from the EU, the Food and Agriculture Organization of the United Nations (FAO), UNEP, and the World Bank**.** This is discussed in further detail below.

3.2.5.1 Survey and qualitative evidence of policy impact

The survey of beneficiaries shows that a majority of surveyed beneficiaries indicated that they disseminated their project results by discussing them with policymakers (60%, see Figure 17).

Figure 17 Beneficiaries' views on the sustainability of their project results



Self-reported project outcomes also indicated multiple interactions with policymakers, in the form of expert testimony or face-to-face advice and consulting. Many (37 records from 12 funded projects) of the 167 workshops previously mentioned in Section 3.2.4.2 were also reported by awardees to include policymakers in their audience. In awardee reports on societal outcomes of projects, 18 out of 65 projects (CRA 2013–2016) reported policy outcomes that have been realized through supported teams' work. Here again, examples provided in the answers included workshops with policymakers, advice, syntheses of evidence conducted for bodies such as the IPCC or IPBES, as well as the creation of new organizations or policy-oriented networks of expertise, and one instance of the creation of a new environmental

protection zone (see the box below).

Policy change and expert policy advice outcomes from awardee reports

The Jordan Water Project (Freshwater 2012) was conducted in close collaboration with the Jordan Ministry of Water and Irrigation (MWI), the Jordan Valley Authority and the Water Authority of Jordan. This collaboration has notably led to a follow-up project (with the MWI endorsing both projects) funded by the Belmont Forum, FUSE (NEXUS 2016). The German project team received support to conduct a series of capacity-building workshops with Jordanian engineers and scientists in a complementary project called JordanCap. This was a key element enabling the project's work to be used by those in Jordan who need it most. When contacted about potential recent outcomes from this project, Professor Steven Gorelick stated:

> "Our development of a country-wide integrated model of Jordan's water system was received with great interest by the Ministry of Water and Irrigation, with whom we maintained excellent cooperation. We have strong indications that they will continue to use the model to help with their long-term planning. They found particularly valuable our evaluation of the economic well-being aspects of freshwater supply and demand, and our model's ability to evaluate policy interventions ranging from tariff increases to new infrastructure… Having the Minister so closely connected to our project was essential to our success. We obtained more data and more expert input than we would have received otherwise. The Minister was amazingly knowledgeable about every aspect of the water sector, from hydrology to operations to policy. I maintain contact with him through the Middle East Water Forum, a group that he has initiated, and I serve on its advisory board. Currently we work directly with the MWI's Secretary General and we have an endorsement from them for our continued work."

Pan-Arctic Options (Arctic 2014) project team members have been active in fostering exchanges between some of the policymakers (from the eight Arctic states) that have signed the recent *Agreement on Enhancing International Arctic Scientific Cooperation*. In an email communication that aimed to identify any long-term policy outcomes that might have been realised after closure of the project, Professor Paul Berkman mentioned that "[o]ne policy outcome is the development of science diplomacy and informed decisionmaking among foreign ministries, as reflected by training and service on advisory boards with national foreign service institutes as well as initiatives and appointment with the United Nations Institute for Training and Research (UNITAR)."

The Nile-NEXUS team's (Mountains 2015) work contributed to the 2017 decision by the Government of Ethiopia to make Choke Mountain a protected zone. The project team declared in its awardee report that "this policy decision results from many years of efforts by stakeholders and researchers. The food-energy-water framing of our Belmont Forum project helped to provide a broad foundation for the closing arguments in establishing the reserve, and analyses pursued in the development of the project contributed to the scientific body of knowledge that demonstrated the importance of the highland forests to the region".

Expert testimony and consulting interactions were conducted as part of Belmont Forum-funded projects for bodies such as Iceland's special envoy on Ocean Affairs; IFWEN investigators advised the mayor and civil servants in the city of Sao Jose dos Campos (Brazil); and members of the Gold Matters project have exchanged with policymakers and local regulators in Burkina Faso and Brazil. They have notably alerted these audiences to the contribution of the mining industries to insecurity in the first country. In Uganda, they have assisted a women-led NGO in devising mining practices to reduce mercury exposure.

3.2.5.2 Frequency of citations toward peer-reviewed publications in policy documentation

Benefiting from the recent emergence of databases that systematically capture and parse policy documents (including white papers, parliamentary and other institutional deliberation transcripts, and legislative texts) made available online, it was possible to examine policy and science advice citations made toward Forum-supported journal publications. Governmental scientists commonly refer to evidence from scientific publications in supporting their arguments or synthesizing available evidence with a view toward decision-making, using references that mirror those found in journal articles and that can now be recorded on a large scale. The Overton database used for this component of the evaluation systematically indexes more than two million documents produced by governmental agencies of all levels – for example, national parliaments and their research services, IGOs, and NGOs such as think tanks and others.

It should be noted that while the Overton database does archive executive and legislative documents such as transcripts of parliamentary sessions, the vast majority of citations toward Forum-funded publications found in the database originated in either IGO reviews of current science or government science advice grey literature. Therefore, the indicator used captures the first steps along the science–policy continuum but cannot be considered as evidence of comprehensive policy change that might have been brought about by project outcomes.

Additionally, the analyses conducted on policy citations were designed with a different approach to those presented so far. Again, we refer to Appendix F for more details on the design of this component of the analysis.

Examples from the Overton records on Forum-funded peer-reviewed publications

Six Forum-supported publications (originating in CRA 2012–2014 projects) were cited in UK Parliamentary Office of Science and Technology POSTnotes. The Office describes its POSTnotes as "resources [that] are for people who read research to inform their opinions or to help them scrutinize legislation and policy decisions."[20] The organization aims to produce "impartial, non-partisan, and peer-reviewed briefings, designed to make scientific research accessible to the UK Parliament."[21]

The Directorate-General for External Policies of the European Parliament published an analysis in 2018 on The Impact of the Common Agricultural Policy on developing countries. The study was requested by the European Parliament's Committee on Development and conducted by Professor Maria Blanco of the Universidad Politécnica de Madrid. It included a citation to work conducted by DEVIL (Food 2013) investigators, "Reducing greenhouse gas emissions in agriculture without compromising food security?"[22]

An example of citation originating in a think tank document is provided by the white paper Calling for Nexus Thinking in Africa's Energy Planning, produced by the Italian Fondazione Eni Enrico Mattei. The cited publication was the SAHEWS (Freshwater 2012) publication "Climate and southern Africa's water–energy–food nexus".[23]

As presented in Table 10, **31.8% of Forum-supported publications resulting from the 2012–2014 CRAs were cited at least once by a policy document** from the Overton database. Notable policy users of Forum-supported research include the UK Parliamentary Office of Science and Technology, the EU, the FAO, UNEP, the World Bank and various think tanks.

Looking at the full Forum-supported publication set, the policy citation rate drops to 22.2%. However, it should be noted that it may take three to four years, or even longer, for the greater share of policy citations to be realized. Therefore, Forum-supported publications from recent CRAs account for this decrease because they had not accumulated citations over a long-enough period at the time of writing these lines.

To assess and interpret the Forum's performance in terms of policy and science advice citations, we compared the results of this analysis in two different (sub-)fields:

- **In comparison A (in Table 10, against BiodivERsA for publications in the subfield of Ecology),** Forum CRA 2012–2014 publications reached a policy citation rate of 36.8%. BiodivERsA papers (which were mostly in this subfield) were cited in a proportion of 29.6%. The

[20] The Parliamentary Office of Science and Technology (n.d.). Understanding research evidence. Retrieved on 05.07.2020 from https://post.parliament.uk/understanding-research-evidence/

[21]he Parliamentary Office of Science and Technology (n.d.). About. Retrieved on 05.072020 from https://post.parliament.uk/about/

[22] Frank, S., et al. (2017*). Reducing greenhouse gas emissions in agriculture without compromising food security?* Environmental Research Letters, 12(10), 105004.

[23] Conway, D., et al. (2015). *Climate and southern Africa's water–energy–food nexus*. Nature Climate Change, 5(9), 837-846.

difference is not statistically significant, but given the restriction in available data, we consider it is safe to conclude that both funders sharing a similar model of funding international and interdisciplinary collaboration leads to similar levels of policy uptake.[24] When looking at comparison B in Table 10, we see that the Forum scores only slightly lower than it does in comparison A. Since it scored markedly above the NSF and NERC in this case, one could hypothesize that the collaborative model of Forum and BiodivERsA favor the policy uptake of research findings relative to more general funding models (ERC being in a separate category focused on research excellence). Notable sources of policy citations to the BiodivERsA publications included the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Naturvårdsverket (Swedish Environmental Protection Agency).

- **In a second comparison (comparison B in Table 10), Forum publications'** policy citation achievements were appraised in relation to those of the ERC, the NSF and the NERC. This comparison was restricted to the Meteorology and Atmospheric Sciences, Environmental Sciences, and Ecology subfields. The share of publications cited in policy documents was 34.2% for Forum, 33.1% for the ERC, 18.6% for the NSF, and 20.6% for the NERC. The small lead of Forum accomplishments over those of the ERC was not statistically significant. The Forum's leads on the NSF and NERC observations, however, are large enough to leave no place for uncertainty (they are statistically significant). The three benchmarking groups mobilized here received notable policy citations by UNEP, the Arctic Council, the World Meteorological Organization (NSF) and the International Union for Conservation of Nature (NERC).

Table 10 Share of Forum publications' (and of benchmarking groups' papers) cited at least once in policy documents, 2009–2019

Group	Total N pubs	N pubs pol. cited	Share pol. cited	Stability intervals	p-value (BF lead)	Notable pol. citations
Descriptive statistics						
Belmont Forum overall	654	145	22.2%			
Belmont Forum, CRAs between 2012–2014	368	117	31.8%			EU; FAO; UNEP; World Bank
Comparison A: Ecology subfield						
Belmont Forum, CRAs between 2012–2014	68	25	36.8%	(25.3%, 48.2%)		as above
BiodivERsA, 2008 call	142	42	29.6%	(22.1%, 37.1%)	0.148	IPBES; Naturvårdsverket
Comparison B: Three main subfields						
Belmont Forum, CRAs between 2012–2014	199	68	34.2%	(27.6%, 40.8%)		as above
European Research Council	199	66	33.1%	(26.6%, 39.7%)	0.412	Arctic Council; EU; UNEP
National Science Foundation, US	199	37	18.6%	(13.2%, 24.0%)	0.000	FAO; UNEP; WMO
Natural Environment Research Council, UK	199	41	20.6%	(15.0%, 26.2%)	0.001	IUCN; UNEP

Note: Main subfields for comparison B: Ecology; Environmental Sciences; Meteorology & Atmospheric Sciences. The number of papers used to compute the share of papers cited in policy documents is lower than the total number of papers for each benchmarking group because papers with no DOI in Scopus could not be queried in Overton (the policy database). P-values are based on a one-tail test for a null hypothesis of Forum smaller than or equal to the benchmarking group.

Source: Prepared by Science-Metrix using the Overton and Scopus (Elsevier) databases

[24] Pinheiro, H.N., Vignola-Gagné, E. & Campbell, D. (2020). *Using Overton policy citations in assessing the uptake of cross-disciplinary research in decision-making*. Science-Metrix. In preparation.

3.2.6 Any other effects

EQ7: Other effects and outcomes: Did the Forum have any other effects (intended or unintended, positive or negative)?

This question looks at any other potential effects that may be linked to the Forum but that might not be captured by the above-mentioned evaluation questions. These effects may be positive or negative, and intended or unintended.

The chapters on this report elaborate on a range of effects, some of which were intended, others that were unintended. These effects are all described in the chapters on effectiveness, efficiency, coherence, added value, and sustainability. The evaluation did not identify any other effects that did not fit in these chapters.

3.2.7 Overall impact of the Forum

EQ8: Overall impact: To what extent has the Forum (in collaboration with all its stakeholders) contributed to the challenge of understanding, mitigating, and adapting to global environmental change?

All in all, the Forum had considerable positive effects, which are likely to contribute to enhancing our understanding of global environmental change, as well as our ability to mitigate and adapt to this change.

3.3 Efficiency and organizational set-up

Key findings on efficiency and organizational set-up

- The Forum's governance structure is practical and functional. It supports the achievement of the objectives, and it divides control and responsibilities in a clear and balanced way.
- With the adoption of the MoU, the member organizations agreed on launching calls on a periodic basis using a model of a joint call and evaluation but coordinated funding, according to organizational rules, with no requirement for funding across borders (but this can be done when their organizational mandate makes this possible, in this case it often allows more inclusive participation, particularly from low and middle income countries. It also describes how members of the Forum interact in the calls. This part of the governance works well and leads to impact. A large majority of surveyed beneficiaries is satisfied or very satisfied with the Forum's processes around the CRAs.
- Where the focus of the MoU is on the calls, the ToR describe (predominantly) the organization and procedures of the Forum outside the calls.
- Here the role of the Secretariat is very important. It provides support to the SC and the co-chairs, organizes the Plenary Meeting, and implements the decisions taken at annual meetings and by the SC. The Secretariat provides support to the CRAs in many ways: in the scoping phase, with ICT (BFgo: helping with the launch, answering questions submitted by proposers), training (e.g. training of researchers in transdisciplinary research) and in monitoring the project progress and implementation (e. g. participating in weekly telecoms, helping with the launch, organizing progress workshops and valorization meetings). The Secretariat interacts through electronic communication means and meets physically at least once a year.
- The total costs for the Secretariat are around 4%–5% of CRA budget.[25] In the experience of Technopolis, program management costs for straightforward national programs are in the range of 7%–10%. Higher percentages (12%–15%) are not unusual for complex programs requiring much coordination. The Forum is such a complex initiative, and the Secretariat is managing quite a bit of the complexity. Although a fair amount of CRA preparation costs and the full costs for project selection and financial administration are with the members, the ratio of execution costs against program funding seems to be low for the Forum. In this respect the Forum Secretariat seems to be operating efficiently.
- The present composition of the Secretariat is, however, not sustainable. The staff is overcharged, and despite the high effectiveness and efficiency as presented above, there is increasing dissatisfaction with (some) members (especially regarding communication). This can be solved by either reducing the tasks of the Secretariat or increasing the staffing of the Secretariat.
- The rules and procedures of the Forum are transparent and inclusive. Despite the high concentration of countries from the Global North as members, the Forum did in fact manage to establish collaboration between the Global North and the Global South. In the co-publication analysis, the Forum scored significantly higher in terms of North–South collaboration than other funders. However, there is still room for improving the involvement of Southern researchers and organizations into the Forum.
- The internal CRA procedures (i.e. scoping of the calls, drafting of proposals, selection of the research proposals) are satisfactory to the members and the beneficiaries. The Forum seems to be an efficient mechanism in terms of coordinating and promoting the proposal processes.

The evaluation of the efficiency of the Forum forms an important component of this study, not only due to the size of the Forum, but also its international nature (staff are based around the globe, many coordination and communication activities take place online) and the various types of stakeholders involved (members, partners, researchers, other project participants, etc.). The term "efficiency" can cover a variety of relationships between effects on the one hand and costs on the other.

This chapter examines to what extent the Forum has carried out its work efficiently in terms of day-to-day operations and the division of tasks, and coordination. In view of the significant growth of the Forum, this chapter also contains an assessment of the Forum's financial and human resources. Lastly, this chapter elaborates on the appropriateness to the way in which the Forum is organized.

To do so, we assess how the Forum is governed, managed, and organized, we look at the day-to-day operations, the efficiency of internal procedures and coordination and the transparency and inclusiveness of external procedures and interactions with stakeholders.

3.3.1 Governance and organizational set-up

EQ9: How appropriate are the governance and organizational set-up of the Forum (especially considering its growing membership)?

3.3.1.1 The Forum's governance structure

The governing documents for the Forum consist of two main documents, namely:

1. The MoU, signed by 10 founding partners from nine countries[26] (in 2012); and
2. The ToR that were drafted in 2015 (an update of these ToR was discussed in 2019, but not implemented). The Forum has no legal status of its own (which creates several problems including submission of contributions, hosting agreements, status of staff, etc.).

With the adoption of the **MoU**, the member organizations agreed on launching calls on a periodic basis using a model of a joint call and evaluation but national funding, according to national rules, with no requirement for funding across borders. It also describes how members of the Forum interact in calls.

This part of the governance seems to work well. Seventeen CRAs have been launched, leading to a large number of projects addressing the Belmont Challenge, and with research results that are of high academic value and large societal relevance and uptake (as described in the previous chapter of this evaluation report). A large majority of surveyed beneficiaries is satisfied or very satisfied with the Forum's processes around the CRAs (between 70% and 80% on different sub-criteria). One area for improvement that we found in interviews and survey, however, was the CRA scoping process. This process is in large part done by funding and resource organizations. The survey results and interviews suggested that the scoping process for new CRAs was often quite limited: only a small number of stakeholders participated in the scoping workshops, and the selection of these stakeholders was often not clear. The move to online scoping that became necessary because of Covid19 seems to have changed this and

[25] Our assessment of these costs exclude the costs for this external evaluation, but include all activities of the Secretariat (including SC and plenary support and communication).

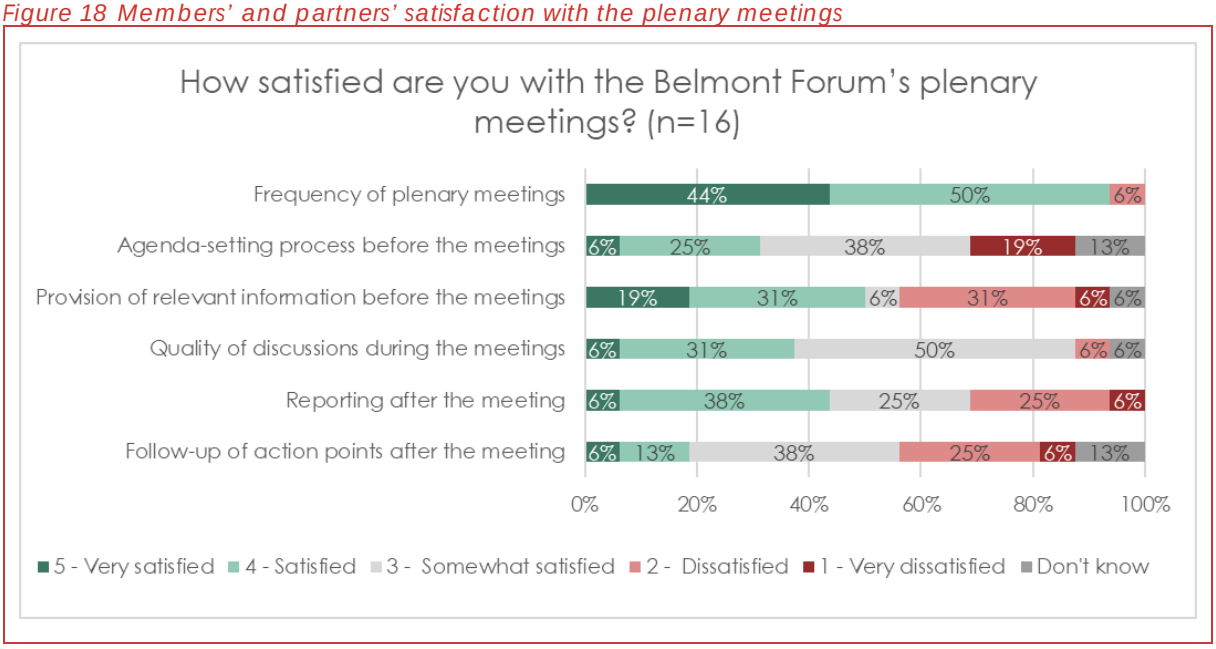
[26] Two of the Belmont Forum's founding members were from the UK.

made increased expert and funder participation possible. Due to the timing of this report, however, we did not receive extensive feedback on the effects of this.

Interview and survey results show a very mixed sense of satisfaction of members with the **ToR** part of the Forum's governance (satisfaction levels ranged from "very satisfied" to "very unsatisfied" here). The sections below provide more detail on the issues that the Forum encounters (related to the two main governance bodies of the Forum, the Plenary Meeting and the SC). The role of the Secretariat will be discussed in chapter 3.3.4 on Human Resources.

3.3.1.2 Plenary Meeting

Figure 18 Members' and partners' satisfaction with the plenary meetings



Source: Technopolis Group, 2020

The Plenary Meeting is the highest authority within the Forum's governance structure and consists of principals from member organizations who can take executive decisions regarding the mobilization of resources to support Forum activities. After 10 years, this meeting has grown from a small club of founding fathers at very high level in their organization with strongly shared views (since they developed these views together) to a much broader group of representatives (not always of the highest level in their organization) with a far greater diversity: some members participate in quite a number of CRAs while others participate in not more than one CRA. The role of partners and observers during the plenary is also not clear to everyone.

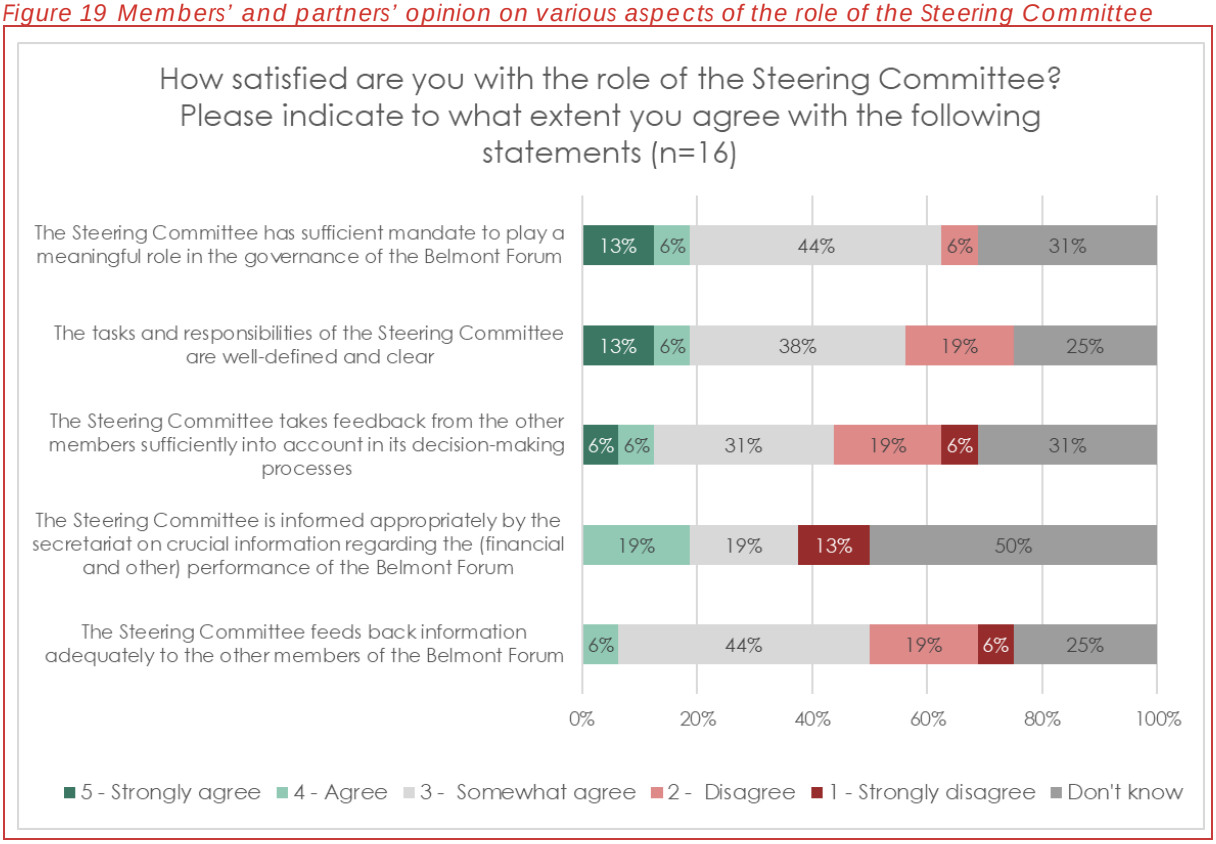
Figure 18 provides an overview of members' and partners' satisfaction with the plenary meetings. The frequency of the annual plenary meetings is considered adequate and the quality of discussions is also seen as at least satisfactory by most respondents. However, a significant minority is (very) dissatisfied with the agenda-setting process. This dissatisfaction does not seem to be caused by the planning[27], so it must be the implementation. The efficiency of the agenda-setting activities should be improved by all involved, so that the representatives in the Forum can get the feedback from their top-level management as well as a specific

[27] In fact, requests for documents and relevant agenda items are circulated in June to allow sufficient time for a September delivery to all members and attendees. This is six weeks prior to the meeting.

mandate to act during the meeting. About one third of respondents are (very) dissatisfied with reporting after the meeting and follow-up of action points after the meeting as well. Communication with members seems to require more attention than it is getting right now.

3.3.1.3 Steering Committee

Figure 19 Members' and partners' opinion on various aspects of the role of the Steering Committee



Source: Technopolis Group, 2020

An important finding that stands out when looking at the opinion of the Forum's members and partners about the role of the SC (Figure 19) is that many (between 25% and 50%) indicated they were not able to answer these questions.[28] This finding shows that communication between SC and members can (and should) improve. If non-SC members of the Forum do not know what exactly the mandate of the SC is (and whether that is sufficient) and what their tasks and responsibilities exactly look like, then this is an important indication that there is a lack of engagement, interaction and communication between the members and SC. Interview results also suggest that the mandate of the SC vis-à-vis the Plenary Meeting is unclear.

3.3.1.4 Other governance aspects

Legal status

The Forum has no legal status of its own. This has the disadvantage that no staff can be appointed (e.g. the Director of the Secretariat has an employment contract with IAI), that no

[28] This can partly be attributed to the fact that only SC members can answer these questions. This applies especially the 4th and 5th question in the figure below, although many respondents are either current or former SC members.

legal contracts can be signed (which might be a disadvantage when attracting funds from non-members), and that there are limitations for some members to transfer membership dues to a non-legal organization (and inability to be a grantee due to its international location). Creating a legal personality for the Forum would have certain advantages, but the practical implications have been limited so far, and creating a legal personality for the Forum would probably require approval from all members at the highest level (sometimes even at cabinet level) which might be difficult to achieve.

Composition of SC and Secretariat

In relation to the composition of the SC, the ToR only addresses a minimum number of four members. The Forum could consider including rules to stimulate diversity in representation here as well. This is also valid for the composition of the Secretariat that now has a very strong (in the eyes of some too dominant) NSF component (provided however without cost for the Forum). At the same time, it is important to recognize that while such composition rules could empower participation by some countries, they might also reduce flexibility (for all positions) to make use of good candidates that show willingness and motivation to perform additional tasks at little or no costs for the Forum.

3.3.2 Transparency and inclusiveness

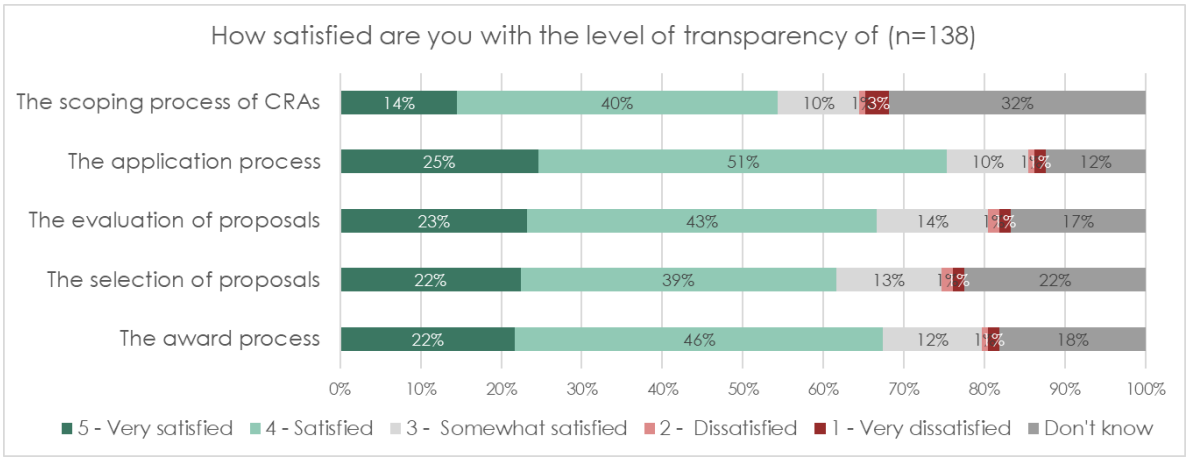
EQ10: Are the rules and procedures of the Forum appropriate (e.g. transparent and inclusive)?

This section elaborates on the Forum's level of transparency and inclusiveness. Whereas transparency mainly relates to the extent to which the Forum's processes and procedures are clear and transparent, the level of inclusiveness relates to the extent to which the Forum is equally accessible and inclusive to various types of stakeholders, including from different sectors and geographical backgrounds. We asked the opinion of beneficiaries.

3.3.2.1 Level of transparency

Most surveyed beneficiaries are satisfied or very satisfied with most aspects of transparency (Figure 20). The respondents seem to be being less aware of the transparency of the CRA scoping process (32% of respondents indicate "don't know", which is not a surprise since this process is largely done by funders to prevent conflict of interest). It should be noted that this is the view of successful applicants only since data of non-successful applicants are confidential and were not disclosed to the evaluators.

Figure 20 Beneficiaries' views on the transparency of CRA development and proposal evaluation



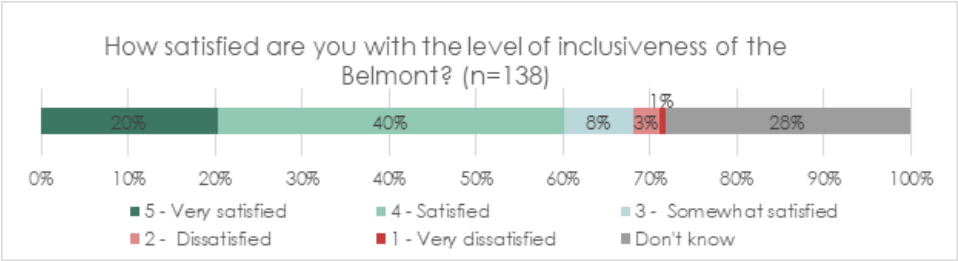
Source: Technopolis Group, 2020

In terms of financial transparency, we did not perform a financial audit of the Forum's finances as part of the evaluation: Technopolis is not an auditing firm and an audit was not part of our evaluation proposal. At present, the IAI provides accounting services under its hosting services. However, it was unable to provide accounting for annual structured financial accounts *(providing similar information as a Profit & Loss account and a Balance Sheet for legal entities),* given the absence of information, receipts, and other required documentation in an organized, detailed and timely manner.

3.3.2.2 Level of inclusiveness

The findings of this evaluation show that beneficiaries of the Forum were largely satisfied with the level of inclusiveness of the Forum (60%).

Figure 21 Beneficiaries' views on the level of inclusiveness of the Forum



Source: Technopolis Group, 2020

An important element in assessment the level of inclusiveness is the involvement of both members and beneficiaries in the activities of the Forum. A review of Forum projects indeed highlights the **concentration of project partners among the Global North**, with those from the Global South being concentrated among a few large member countries, such as Brazil (São Paolo), India, and South Africa that were funding partners for the calls. The low participation is largely caused by a lack of funding for research in most other countries of the Global South, particularly low-income countries.

Members and partners of the Forum recognized this. In the survey of members and partners, two thirds of the members felt that the geographical spread of the Forum's members was too limited and needed to be diversified, especially to increase participation from countries and researchers from the Global South.

Even though there is considered to be room for improvement, the Forum has provided an **opportunity for participation by researchers from and in the Global South in transnational global environmental change research, as both Principal Investigators and other personnel in projects**. We have not identified another research funding initiative which has provided this. Two thirds of beneficiaries from the Global South who responded to the survey indicate that their project involved co-design or co-creation to a large extent or more (see Figure 22 and the case study on North–South collaboration in Appendix E).

Figure 22 Beneficiaries' perspectives on co-design and co-creation in projects

Source: Technopolis Group, 2020

In addition (as detailed in Appendix E), the Forum may have contributed to **increased joint publishing opportunities for participating researchers from the Global South**. The bibliometric analysis shows that ICRs including participation by the Global South are over 40%. This is the same level achieved by publications funded by both Chinese and Brazilian funding agencies included in the analysis (data not shown). It seems unlikely that such a level would be seen among the participating researchers from the Global South in the absence of the funding and collaboration mechanisms provided by the Forum. Lead or corresponding authors from the Global South account for 31% of Forum publications. Although somewhat lower than the proportion for global publications in similar thematic areas, that comparison is less insightful given the predominance of non-international publications from some large countries of the Global South.

The Forum has thus achieved substantial results in terms of inclusion among its current members from the Global South. Broadening inclusion of the Global South requires more members and a means, or new funding models, to support their participation, particularly for researchers in low-income countries. A principal constraint is that funding agencies from the Global South generally have far fewer resources than those in the Global North, especially in the case of low-income countries. The model of the Forum is based on members and partners mobilizing and contributing their respective resources. This model entails challenges to promote greater inclusion of researchers from the Global South. The interviews and survey of members of the Forum indicate the desire to increase the share of members (and beneficiaries) of the Forum even further. To do so, one would need to consider a new or adjusted funding model that would support participation of countries from the Global South, particularly low-income countries. Such support could come from development agencies and has taken place in the

past (SIDA support to participation of low- and middle-income countries in the Oceans CRA, UKRI in CEH, and Future Earth in Pathways).

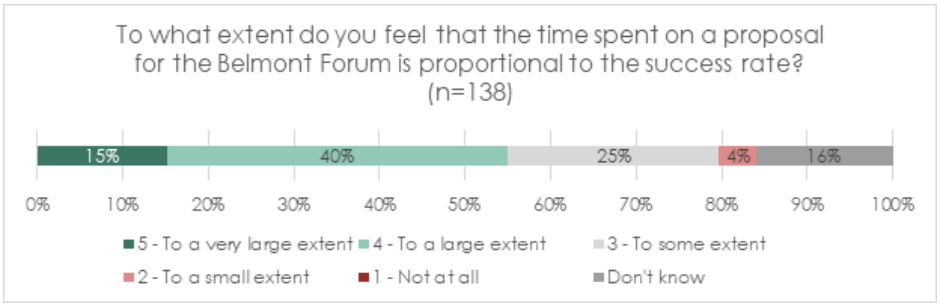
3.3.3 *Efficiency of internal (CRA) procedures*

EQ11: To what extent has the Forum carried out its work efficiently (e.g. financial and human resources, internal procedures, etc.)?

Evaluation question 11 relates to the efficiency of the Forum's main procedures.

Regarding the proposal submission procedure, over half of the surveyed beneficiaries (55%) are satisfied with the ratio between the efforts needed for writing proposals and the success rate, as shown in Figure 23.

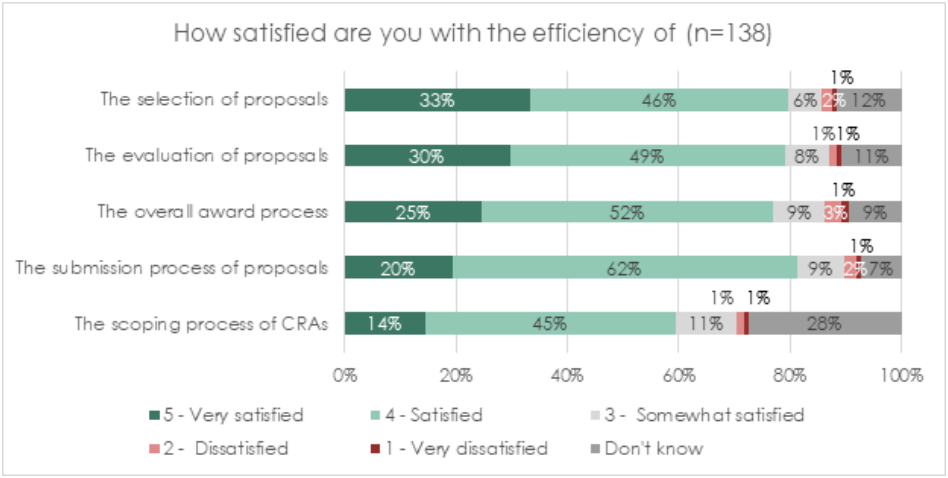
Figure 23 Beneficiaries' views on the effort required for proposal writing



Source: Technopolis Group, 2020

Beneficiaries also seem mostly satisfied with the efficiency of other proposal-related processes of the Forum. Between 55% and 82% was either "satisfied" or "very satisfied" with the various processes. Very few of the surveyed beneficiaries are explicitly negative of these processes, as shown in Figure 24.

Figure 24 Beneficiaries' views on the efficiency aspects of CRA development and proposal evaluation



Source: Technopolis Group, 2020

The findings from the survey and interviews with members and partners show similar results. Even though international cooperation is complex, the Forum seems to be an efficient mechanism in terms of coordinating and promoting the proposal processes.

Opportunities for further improving the efficiency for beneficiaries seem to lie predominantly in the alignment of reporting requirements (to the national funder) with Forum monitoring demands. This either requires a stronger centralization of reporting (meaning less autonomy for the members and more work for the Secretariat) or accepting reduced insight in what is happening in the Forum.

For some members (especially in Europe, but also in other countries where research councils are more and more losing freedom in their allocation of funds and more and more funding is in focused programs), synchronization of their own funding procedures with Forum CRA procedures is becoming increasingly difficult.

3.3.4 Human resources

EQ12a: Are the Forum's human resources appropriate for the work entrusted to it?

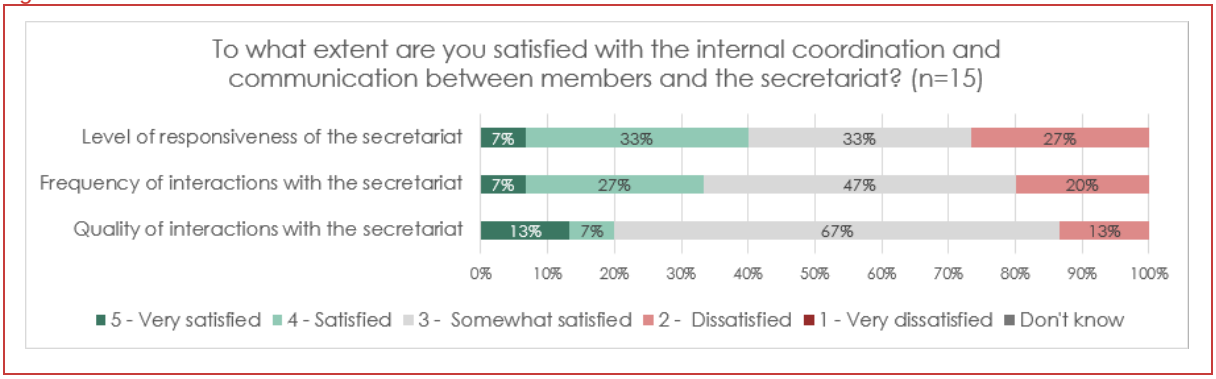
The ToR of the Forum describe the tasks of the Secretariat only in broad terms. No specific additional descriptions of the priorities for the Secretariat are available. It is, however, clear that the activities of the Secretariat have increased over time, due to increase in number of members and number of CRAs, as well as because of a gradual but continuous increase in expectations from the members. When the CRAs started, there was a need felt for consistency in CRA implementation and communication, central monitoring of progress and impacts, working with the TPO to develop the valorization activities, etc. When membership numbers started to grow, a need for formal onboarding processes for new members and more consistent communication with members started to arise.

Currently, the Secretariat provides support to the SC and the co-chairs, organizes the Plenary Meeting, and implements the decisions taken at annual Forum meetings and by the SC. The Secretariat also provides support to the CRAs in many ways: in the scoping phase, with ICT (BFgo: helping with the launch, answering questions submitted by proposers), training (e.g. training of researchers in transdisciplinary research) and in monitoring the project progress and implementation (e. g. participating in weekly telecoms, helping with the launch, organizing progress workshops and valorization meetings).

This all piled up at the Secretariat without really increasing capacity at the Secretariat, and without making explicit prioritizations for the Secretariat. **It is generally acknowledged that the Secretariat lacks capacity in relation to all the tasks they perform. The present composition of the Secretariat is not sustainable** (with one director, two full-time AAAS fellows (an in-kind contribution by the NSF) and four part-time (0.2 FTE each) staff (as in-kind contributions of other members)(see also below, 3.3.5)).

Despite the large number of Secretariat activities as presented above, the evaluation show that while most members indicate to be at least "somewhat satisfied" with the functioning of the Forum, there is also a number of members indicate to be dissatisfied with the Secretariat (Figure 25). The number of dissatisfied members ranges between 13% and 27%, depending on the specific issue (a similar level of dissatisfaction as shown with the SC). There seems to be room for improvement in the internal coordination and communication between members, the SC, and Secretariat.

Figure 25 Members' satisfaction with the Secretariat



Source: Technopolis Group, 2020

The interviews confirm an increasing issue over the past years in the communication between at least some members and the Secretariat. Complaints from these members relate to the follow-up of decisions taken during plenary meetings, the transparency, and timeliness of information (for example on the Forum's finances or activities). At the same time, it was indicated that members did not always respond to requests to read emails or documents, making it difficult for the Secretariat to meet agreed timelines.

While it is not up to the evaluation team to make final judgments on each of these individual incidents, it is clear from the findings that **the relationship between some members and the Secretariat of the Forum has been negatively affected**. It can be asked here, whether the dissatisfaction is directed correctly to the Secretariat, or should be directed toward the SC, but it needs explicit attention going forward.

3.3.5 Financial resources

EQ12b: Are the Forum's financial resources appropriate for the work entrusted to it?

In the past years members and partners generated annually more than €20m per year to start new CRAs. This is only a very small part of all global changes research in the world.

The Forum currently has 19 dues-paying members. Contributions are €20,000 per member per annum. In addition, the Forum has nine in-kind contributing members. The table below provides an overview (in USD) of the income and expenditures of the Forum since June 2018 (when the Secretariat was hosted by IAI). It is based on the overview as provided by the Secretariat to the Forum SC in early 2020.

Table 11 Balance Belmont Forum 2018–2020

	2018 (Jun-Dec)	2019	2020 (projected)
Starting capital	***$121.760***	***$203.774***	***$213.203***
Income (members)	$159.918	$461.331	$458.950
Other income	$86.785	$0	$0
Total income	***$246.703***	***$461.331***	***$458.950***
Salaries & benefits	$97.980	$187.560	$189.516
Travel	$29.662	$83.036	$20.000
Other direct expenses	$22.655	$98.267	$236.985

	2018 (Jun-Dec)	2019	2020 (projected)
Overheads (IAI, 18%/y)[29]	$14.392	$83.039	$82.611
Total expenditure	***$164.689***	***$451.902***	***$529.112***
Balance (P&L)	**$82.014**	**$9.429**	**-$70.162**
End of year remainder	$203.774	$213.203	$143.041

Source: Technopolis analysis of Belmont Forum (2020) Belmont Forum Budget 19 April 2020[30]

The table shows that besides membership fees there has been no other income over the past two years (and in 2018 this was a transfer from ANR that previously hosted the Secretariat). Large expenditures, besides the salary costs, include the evaluation costs (most of other direct expenses in 2020), the amount of travel in 2019 (to visit members, potential members and Forum organized meetings including 5 CRA kick-offs), the contributions to the Forum's 10 years celebration, and the SRI2020 conference in 2019 and 2020.

The table does not show the in-kind contributions from nine members, the additional NSF contribution of at present two AAAS Science and Technology Policy Fellows at full-time or near-full-time equivalent to support the Forum and the NSF contribution to the maintenance of the website, GSuite, and the BFgo ingress and reporting system, which averages between USD 300,000–400,000 per year.

The average participation per funder in a CRA is around €1.2m (with very large variation, between €60k and €4m (excluding EC participation in Biodiversity II) for generally a 4- to 5-year program). In comparison with this, the contribution of individual members to the Forum is €20k per year, which is high for members with small participations in only one or two CRAs, but low for members with regular large participations in CRAs (but that also have significant CRA handling costs themselves).

Total costs for the Secretariat (excluding costs for this external evaluation, but including all activities of the Secretariat, including SC and plenary support and communication), are around 4%–5% of CRA budget. In the experience of Technopolis, program management costs for straightforward national programs are in the range of 7%–10%. Higher percentages (12%–15%) are not unusual for complex programs requiring much coordination. The Forum is such a complex initiative, and the Secretariat is managing quite a bit of the complexity. Although a fair amount of CRA preparation costs and the full costs for project selection and financial administration are with the members, the ratio of execution costs against program funding seems to be low for the Forum, or even too low.

[29] The overhead cited in the table reflects expenditure but does not reflect the total services and benefits provided by the IAI in the hosting of the Secretariat, including diplomatic privileges.

[30] The most recent versioning of this table would have zeroed out travel for 2020 due to COVID quarantine and captured the received dues since April (which would also adjust the overhead).

3.4 Coherence

Key findings on coherence

- The Forum's thematic prioritization through the CRAs as well as the projects funded by the Forum are coherent with each other. The CRAs (and related projects) are well-chosen and each have a clear link to the Forum's overall mission and objectives. While some CRAs are quite closely related or sometimes even overlap, we do not consider this problematic for the overall coherence of the Forum. This may help to understand certain subject matters in more depth, and to look at similar issues from different angles.
- Regarding the alignment of the Forum with other international programs, it is quite a unique initiative being fully international, focused on transdisciplinary research, and global environmental change. However, there are some overlaps with existing programs or organizations as reported by the members. There is a need for increased external engagement with funding and donors' organizations (aid agencies, philanthropic foundations, etc.) and communication, particularly in the scoping phase of the CRAs.

This chapter on the coherence of the Forum assesses the "internal coherence" of the Forum (i.e. the extent to which the various activities of the Forum are well-aligned and complementary to each other) as well as the "external coherence" of the Forum (i.e. the extent to which the Forum's objectives and activities complement those of other relevant initiatives in the landscape of international change research).

3.4.1 Internal: Coherence of CRAs and funded projects

EQ13: Are the CRAs and funded projects coherent with each other?

Evaluation question 13 assesses the extent to which the Forum's thematic prioritization through the CRAs as well as the projects funded by the Forum are coherent with each other.

Before answering this question, it is important to provide some context and to recapitulate some important conditions under which the Forum operates. First, as discussed in chapter 2.2, the mission and objectives of the Forum are very broad. Second, the resources available to the Forum to fund projects are not unlimited. Third, the identification and selection of CRAs are typically done bottom up. This means that members of the Forum (and Future Earth, for one theme each year) take the initiative to propose topics for CRAs. The proposal is discussed at a plenary meeting and voted on to go into scoping. Status updates, including documentation, are shared with subsequent plenaries. Each CRA must be supported by at least three members to go ahead.

Considering these conditions, it would be unrealistic to expect from the Forum to address its broad mission and objectives in a complete and comprehensive manner, covering all aspects related to understanding, mitigating, and adapting to global environmental change. The issue of global environmental change is simply too broad and big for that. In fact, we would argue that there is no harm in the fact that projects might sometimes address similar subfields or topics. This can even help to understand certain subject matters in more depth, to look at similar issues from different angles, and potentially even to test and weigh project results against each other. Moreover, the portfolio of CRAs and funded projects is bound to have certain gaps in thematic fields, geographical areas covered, etc.

Therefore, what should be assessed and monitored very closely are not the gaps or overlaps between activities, but rather (1) whether these gaps and overlaps do not come from any systematic biases, (2) whether the overall portfolio does not contain any overlaps where

projects might hinder each other, and (3) whether there are not unjustifiable amounts of money spent on topics that are considered either too narrow or irrelevant for the Forum's overall mission.

The first point, on systematic biases in the overall portfolio of CRAs and projects, is discussed in more detail in 3.3.2 (evaluation question 10). Here, it was found that **there is room for improvement in terms of enhancing the involvement of stakeholders from the Global South.**

Regarding the second and third points, the evaluation findings suggest that there are no significant issues. Based on our desk research as well as the findings from the interviews with members, partners, and external stakeholders, **the CRAs (and related projects) funded by the Forum are well-chosen and each have a clear link to the Forum's overall mission and objectives**. Members and partners of the Forum argued that while some CRAs are quite closely related or sometimes even overlap, this was not considered problematic for the overall coherence of the Forum. **While the survey of members and partners showed that not everyone was fully satisfied with the scoping workshops and plenary meetings on the proposed topics, the interviews in turn reveal that this has not led to problematic inconsistencies in the Forum's activities.**

3.4.2 External: coherence with other initiatives

EQ14: Are the activities of the Forum coherent with other initiatives in the context of environmental change?

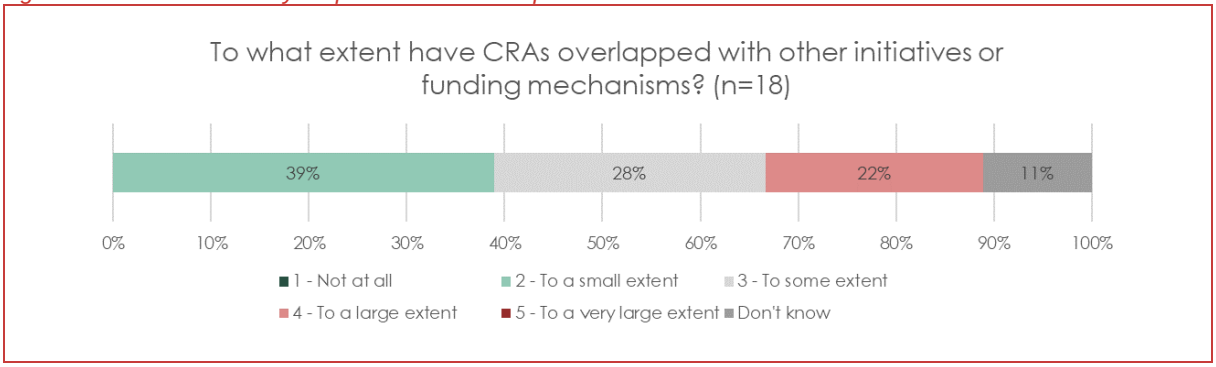
This evaluation question assesses the extent to which the objectives and activities of the Forum are aligned with other similar initiatives in the area of global environmental change research. Given the large number of programs, forums, institutes, and other initiatives that one could potentially look at in the context of this question (at international and national levels), we limited this evaluation to the comparison of a few key international initiatives, in particular, Future Earth, the International Science Council, the Global Research Council, as well as some regional funding mechanisms, such as EC Joint Programming Initiatives. The findings presented in this section were mainly based on the in-depth case study on the coherence of the Forum (which also included findings from desk study, the interviews, and online surveys).

It seems clear from this evaluation that the Forum is quite unique, with no comparable counterparts. It is the only initiative, with a global participation, dedicated to **international, transdisciplinary research** on **global environmental change** (the IAI has a similar mandate but with regional participation). There is a range of initiatives to enhance scientific cooperation and coordination at an international level, such as the International Science Council (ISC) and the Global Research Council. There may be specific attention within these bodies to global environmental change, but they do not operate as funding mechanisms, as the Forum does. There are some regional funding mechanisms, such as the Joint Programming Initiatives of the EC, which also fund interdisciplinary research. However, the focus of the Forum on transdisciplinary research and its broader geographic scale make it generally complementary to such initiatives. Thus, in terms of its mission and the activities it undertakes, the Forum is broadly complementary to other initiatives.

Although the Forum is unique in its mission, scope and membership, various overlaps with other initiatives have been reported by respondents to the survey (see Figure 26). These refer to specific CRAs, and in general, it seems reasonable and to be expected that there is some overlap (overlap to a small extent or to some extent) between CRAs and some other research and funding initiatives.

Partial overlap might only occur on some dimensions. For example, other initiatives are funding research on global environmental change, but this might not be as transdisciplinary as that funded by the Forum. Partial overlap can occur in terms of geographic scope. About a quarter (22%) of members and partners surveyed indicated that the CRAs overlap "to a large extent". It needs to be recognized that the Forum is designed in a way to leverage and couple with national and regional initiatives. Overall, it is difficult to interpret whether the perception among members is generally positive concerning overlap. This suggests either additional procedures or efforts to **refine the scoping process of the CRAs** to ensure stronger and more apparent complementarity with other initiatives.

Figure 26 Member survey responses on overlap with other initiatives



Source: Technopolis Group, 2020

The complementarity of the Forum with **Future Earth** is a complex topic. In many ways, the two initiatives complement each other in terms of their function and activities. For example, Future Earth provides a platform and has built a broad community but is itself not a funding mechanism. This relationship could be more strongly and clearly articulated and communicated to all stakeholders, particularly those who are only partly engaged in these communities (for example, policy advisors and other users of research outputs) as well as researchers who may be (potential) beneficiaries but are not involved with coordination at the international level at which these organizations operate.

Suggestions to improve external coherence include **increased external engagement and communication**, which implies additional demands on the Secretariat. These efforts could target development aid agencies and philanthropic foundations, both of which are funding some similar research efforts, regional scientific hubs, and private organizations and innovation actors. Efforts in this regard are already being made by the Secretariat, which has engaged with various organizations, including the GEF, IDRC, the World Bank, the Welcome Trust, IGES (among others). For example, one suggestion (made in an interview) might be to develop more of a global forum of funders of global environmental change research.

3.5 Added value

Key findings on added value

- The Forum offers clear added value within the global landscape of science funding and a unique combination of project-level strengths and achievements. It fosters multilateral collaboration among funding agencies and among researchers (including North–South collaboration).
- For researchers, it provides access to research funding, and particularly to support transdisciplinary research projects. The bibliometric analysis shows that the Forum is leading worldwide in terms of interdisciplinarity and multidisciplinarity. Bibliometric indicators for transdisciplinarity are still being developed. The Forum is working on these indicators, together with other supporters of transdisciplinary research.
- Turning to indicators of broad dissemination and societal uptake of findings, the Forum's publications displayed a high level of impact within science advice and governmental research documents. The Forum was trailed closely only by the ERC.

"Added value" refers to the value that the Forum brings, that could not be brought (to the same extent) by other initiatives in the field. Here, one can think of the leverage of investments by national funders, the international community building and networking effects, the enhancement of interdisciplinary knowledge, etc.

EQ15: What is the added value of the Forum (compared to other initiatives at various governance levels)?

To answer this evaluation question (at least on one dimension), a bibliometric comparison to other funders in the field of global environmental change research was conducted to assess the Forum's added values compared to similar initiatives (linked to the question of external coherence). We also addressed the topic of added value through the interviews with the Forum's stakeholders, and the survey with beneficiaries to better understand how they perceive the added value(s) of the Forum (and their relative importance).

The findings from this evaluation show that the **Forum offers clear added value within the global landscape of science funding** and a unique combination of project-level strengths and achievements. The Forum's funding has managed to simultaneously foster:

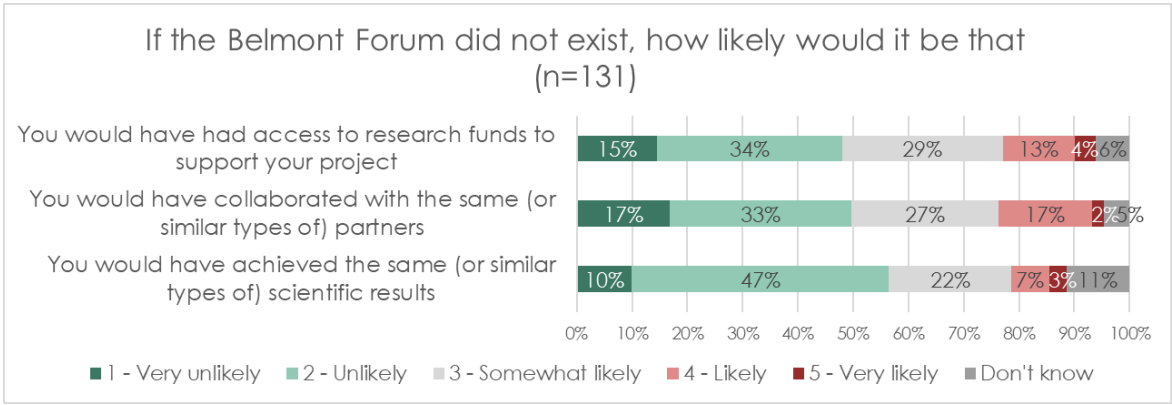
1. Multilateral collaboration among funding agencies as well as among researchers (incl. North–South collaboration);
2. Better access to research funding for researchers and other stakeholders in the field of global environmental change;
3. Research excellence of the highest order and multi- and interdisciplinarity in the production of peer-reviewed publications; and
4. Strong uptake in science advice and research for policymaking.

First, the **multilateral cooperation and coordination of research funds at global scale** was one of the main added values of the Forum. **Members indicated that the Forum lowers barriers to collaboration between funding agencies**. It has led to better contacts between funding agencies, a better understanding of each other's working methods and procedures, the sharing of best practices, and most importantly, more discussion, reconsideration, and, as a result, better alignment of activities. Chapter 3.2.2.2 contains more information on this topic.

The Forum also creates new opportunities for international scientific collaboration. As shown in Figure 27, substantially less of surveyed beneficiaries thought it likely that they would have collaborated with the same or similar types of partners (19%) than those who thought that to be (very) unlikely (50%). Moreover, the Forum clearly led to more collaboration between researchers from the Global North and the Global South, especially when compared to other funders in the field.

Second, the Forum did not only lead to more collaboration between funding agencies, it also led to **better access to research funding** for researchers and other stakeholders in the field of global environmental change. As shown in the figure below, 49% of surveyed beneficiaries considered it (very) unlikely that they would have had access to research funds without support by the Forum. Only 17% of the surveyed beneficiaries considered this a likely or very likely scenario.

Figure 27 Beneficiaries' views on the added value of the Forum



Source: Technopolis Group, 2020

Third, the **transdisciplinary character of the research** funded by the Forum constituted another very important added value. The bibliometric analysis presented in chapter 3.2.2.3 shows this.

Fourth, in terms of uptake of findings of projects funded by the Forum within the **scientific community**, Forum publications also scored extremely well (see section 3.2.3).

Finally, turning to indicators of **broad dissemination and societal uptake** of findings, Forum publications displayed a high level of impact within regulatory science and governmental research documents.

3.6 Sustainability

Key findings on sustainability

- It was not possible and not in the scope of this evaluation to conduct an in-depth review of the research projects that would have permitted to precisely appraise the sustainability of their results. However, members and partners of the Forum provided anecdotal evidence of sustainability. Researchers, in a large portion, have disseminated their project results by combining traditional channels (scientific publications and conferences) with discussions with policymakers (60% of respondent) and non-scientific publications (47%). The fact that the dissemination of project results to various target audiences was largely effective is a promising result in terms of sustainability.
- Regarding the sustainability of the Forum as an organization, there are two challenges to consider and to address in a relatively short term. First, the most pressing issue is the staffing and financing of the Secretariat. Second, there is an increasing divergence among the members in terms of participation and commitment.

This last chapter on sustainability looks at two different aspects. First, we analyze the extent to which the effects and results of projects funded by the Forum are sustainable in the future (i.e. whether they have a lasting effect even beyond the duration of the project). Second, we analyze the extent to which the Forum as an organization is sustainable (i.e. the likelihood that it will survive going forward).

3.6.1 Sustainability of project effects and results

EQ16: To what extent are the effects and results of (projects funded by) the Forum sustainable in the future?

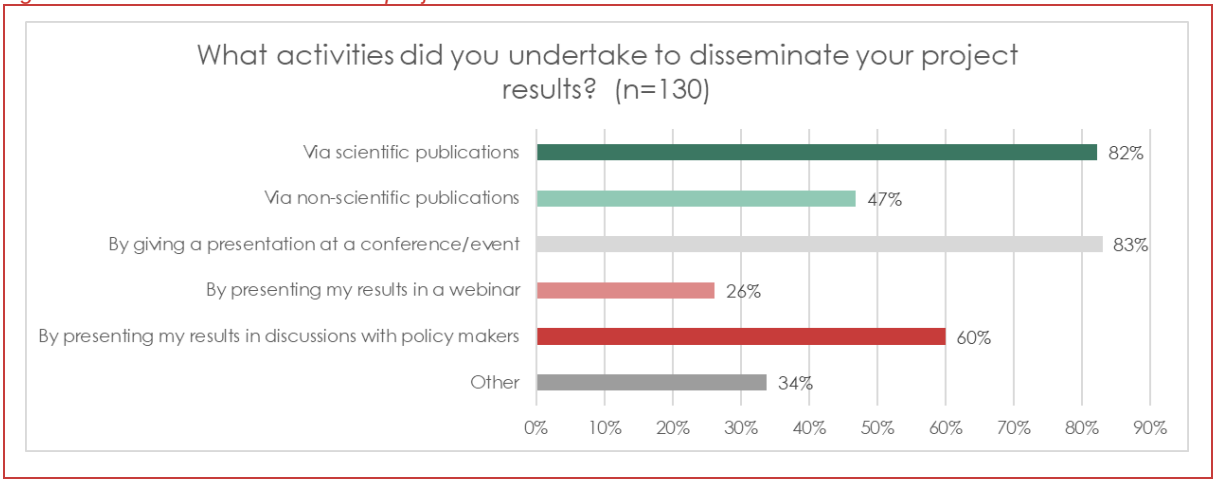
Evaluation question 16 assesses the extent to which the Forum's funded projects have yielded results that are sustainable in the future. For example, it might be that they have led to longer-term collaborations or outputs that can be used and re-used by various actors in further research or in policy, and the like.

To start with, it is important to note that it was not possible within this evaluation to conduct an in-depth review of the results (and their effects) of each project funded by the Forum. Many of the projects are ongoing, while most effects (e.g. policy impacts) will only materialize at the end or after the end of the project. Moreover, the current monitoring system of project results does not allow for a systematic assessment of results, especially not of the earlier CRAs.

Therefore, the answer to this evaluation question mainly relies on a combination of the bibliometric and altmetric analysis on impact, as well as members', partners', and beneficiaries' assessments on the sustainability of the project results. While none of these stakeholders have a full overview of all funded projects, their views and opinions together do provide some indication of the general sustainability of project results in the future.

The findings of the survey of beneficiaries show that a large majority disseminated their project results via scientific publications (82%) and via conferences and events (83%). What is interesting from Figure 28 below is that a large portion (60%) also disseminated their project results via discussions with policymakers. Close to half of the surveyed beneficiaries used non-scientific publications as a channel to disseminate their project results (47%), as shown in the figure below.

Figure 28 Beneficiaries' views on project dissemination



Source: Technopolis Group, 2020

The findings of chapters 3.2.3 on scientific impact, chapter 3.2.4 on dissemination of research results and chapter 3.2.5 on policy impact all showed that the dissemination of project results has been quite effective, as outputs and results funded by the Forum were discussed and taken up by the scientific community, policymakers, and in some cases also by wider audiences. It was found, for example, a great proportion of BF-supported scientific publications were cited in policy documents, indicating that the research had been useful input for decision-making and argumentation by a range of governmental agencies, IGOs and think tanks (even though this may not have led to real impact on policies). Moreover, the Forum appeared to have been effective in disseminating knowledge and fostering discussion of supported research beyond academic circles.

While the longer-term use and accessibility of results after the end of the duration of most projects remain to be seen, **the fact that the dissemination of project results to various target audiences was largely effective is a promising result in terms of sustainability**.

The interviews with members and partners of the Forum also indicated that there was at least some ad hoc evidence of sustainability. A couple of interviewees did mention that they were familiar with some individual projects that had finished, and that did well in terms of impact and sustainability. When asked about ways to ensure sustainability of research results in the future, members and partners of the Forum suggested that the Forum could identify champions and/or showcase best practice examples, etc.

3.6.2 Sustainability of the Forum

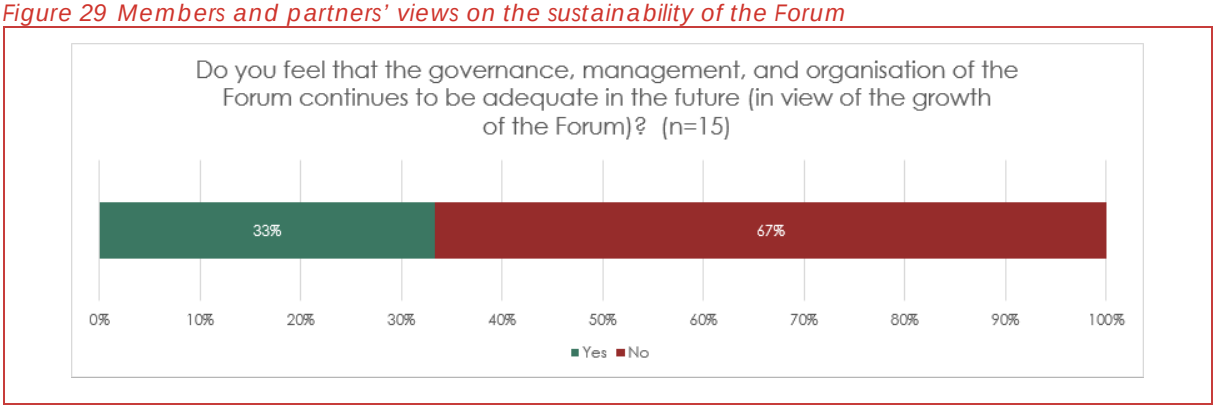
EQ17: To what extent is the Forum sustainable in the future?

This last evaluation question looks at a higher level at the sustainability of the Forum itself. A crucial consideration in answering this question is *whether* and *to what extent* the Forum is likely to be supported by its current and potential new members and partners and has secured already funding for the future.

The MoU between the Forum's member organizations is valid for five years after the last signature of a new CRA partner from a call is approved. This means, in theory at least, that the Forum can continue forever without renewing the explicit consent of the members, provided it continues to attract new CRA partners.

However, two thirds of the surveyed members and partners (67%) feel that the present governance, management, and organization of the Forum are inadequate in the future (see Figure 29). The interviews also reveal that the current budget for the Secretariat, is not considered to be adequate for the Forum going forward at the present ambition level.

Figure 29 Members and partners' views on the sustainability of the Forum



Source: Technopolis Group, 2020

Assembling all opinions about the Forum's organization and governance gathered during this evaluation, there seem to be several interrelated challenges for the Forum's governance:

- **High ambitions with limited capacity:** The demand for more centralization of certain Forum functions has grown significantly over time. Several activities, including project monitoring, training of project leaders, communication, etc. have been improved by centralization of these activities at the Secretariat. Development and technical maintenance costs (USD 300,000–400,000 per year) of these activities to address transnational partner needs, consistency for intercomparison, varying familiarity with open data/access and transdisciplinarity, and the capture of the impact/added value have been borne by NSF, but require also an increasing personnel capacity at the Secretariat. Even though the NSF is supporting that as well (e.g. through the two AAA-fellows), a stable situation requires a readiness to contribute to these functions by all partners.

  To bring the Secretariat at the level of 5–6 FTE, around €800k per year (in cash or in kind) would be necessary to cover costs. To afford this, an increase of membership fees and/or increase in number of members (or other financial contributors) are necessary. Differentiation of contribution (higher fees for "richer" organizations, or fees depending on the participation in/contribution to CRAs) might be an option. Also increasing in-kind contributions to e.g. 0.4 FTE/organization is possible. Finding new members/contributors is not easy and will require large efforts from the Secretariat (or members) to realize. The alternative however is a serious reduction of tasks for the Secretariat, which would either mean a serious reduction in the ambition levels of the Forum and/or a transfer of tasks presently done by the Secretariat to the member organizations.

  The choice of any of these scenarios (or a combination of scenarios) for matching funding and Secretariat costs/activities requires serious discussions and decisions before the end of 2020.

- The second longer-term issue is the **increasing divergence in participation and commitment of members**. Coordinating tasks are done by a limited group of members. There are a couple of members that were very actively participating in the early years of the Forum, but that have now reduced their activities. To some extent this may be caused by budget restrictions for these members, but discussions at the 10th anniversary, and comments

received in survey and interviews, suggest that there is a clear need to take a step back and to reflect on the following questions in order to **(re)create strategic alignment of members and secure the future of the Forum**:

- What objectives should the Forum pursue and how can it align with the SDGs and other international policy initiatives? Should there be more focus, or less?
- What more can be done to include more member organizations and more researchers from the Global South?
- What more can be done to encourage the dissemination and use of research results?

4 Recommendations

Recommendation 1: This evaluation shows that the Forum is relevant and achieving (very) good results in an efficient manner. The Forum should consider the results of this evaluation as an encouragement to build on what is already achieved and continue to initiate and support international transdisciplinary research providing knowledge for understanding, mitigating and adapting to global environmental change.

Recommendation 2: The growth in members and activities means that the Forum is entering a maturity stage. At this point there is a need to discuss whether all members are still in agreement with the Forum's mission and able to ensure a common agreement of the appropriate way forward. The following questions should be considered to (re)create strategic alignment of members:

- What objectives should the Forum pursue? Should there be more focus, or less?
- What more can be done to include more member organizations and more researchers from the Global South?
- What more can be done to encourage the use of research results from the CRAs?

The common agreement resulting from the discussion should be operationalized and reflected in the Forum's mission, objectives, and in the formulation of CRAs.

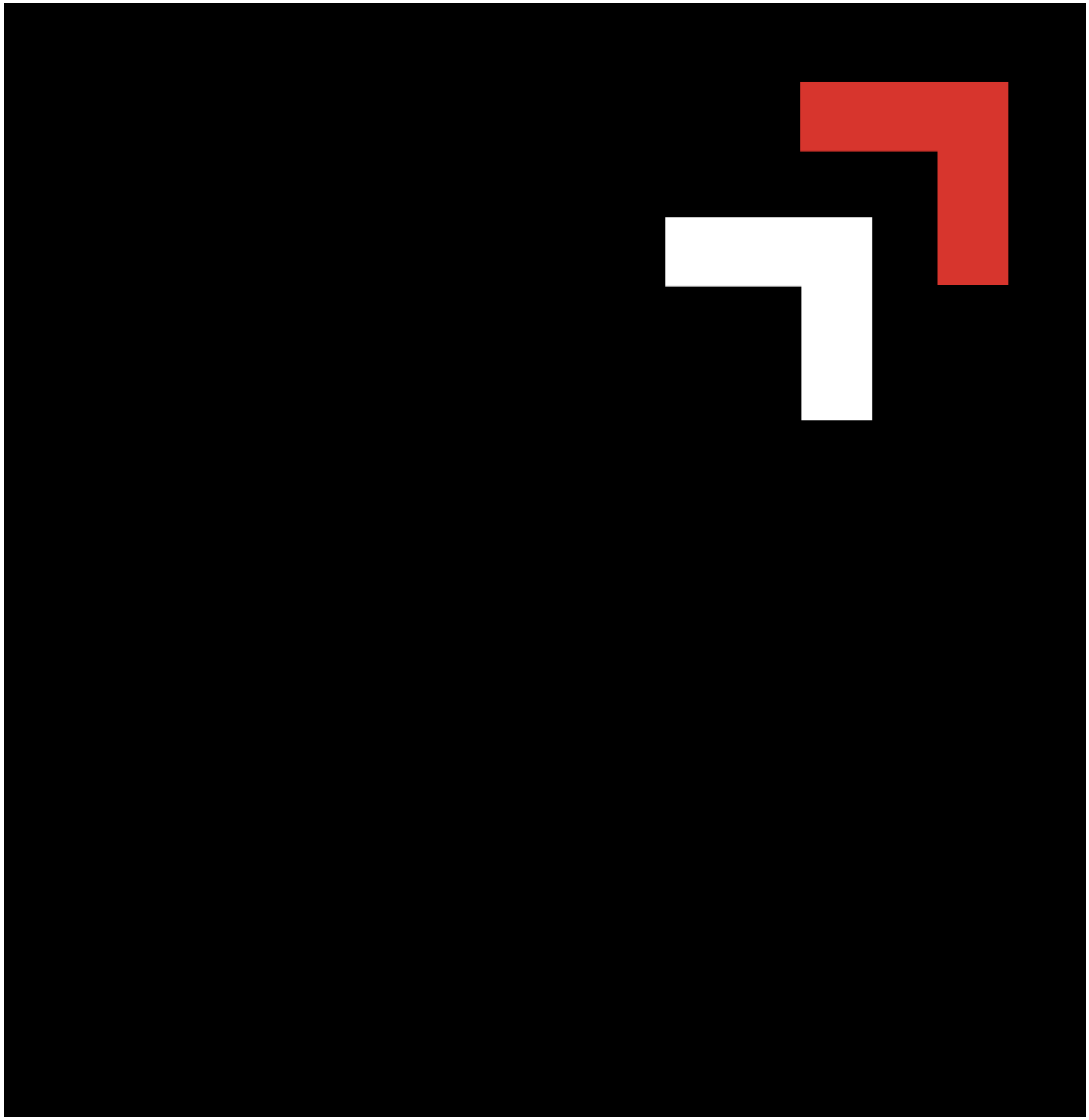
Recommendation 3: The growth in members and activities the demands on the Secretariat have increased, and the Secretariat is understaffed to meet up to the ambitions/expectations. We suggest increasing the capacity of the Secretariat to 5–6 FTE staff members.

Recommendation 4: To bring the Secretariat at the level of 5–6 FTE, around €800k per year (in cash or in kind) would be necessary to cover costs. To afford this increase of membership fees and/or increase in number of members (or other financial contributors) are necessary. Differentiation of contribution (higher fees for "richer" organizations, or fees depending on the participation in/contribution to CRAs) might be an option. Also increasing in-kind contributions to e.g. 0.4 FTE/organization is possible. Finding new members/contributors is not easy and will require large efforts from the Secretariat (or Members) to realize. However, the alternative is a serious reduction of tasks for the Secretariat, which would either mean a serious reduction in the ambition levels of the Forum and/or a transfer of tasks presently done by the Secretariat to the member organizations.

Recommendation 5: In all scenarios for solving present understaffing of the Secretariat, more explicit prioritization of activities of the Secretariat is needed. We suggest addressing this with a more explicit annual planning cycle, with short, to the point, annual work plans with clear priorities (so that activities of the Secretariat match capacity), annual reporting to the Plenary Meeting, empowering of the Secretariat director to implement these annual plans (including more explicit budget responsibility), financial and progress reporting in every SC meeting, and an active interaction between SC/co-chairs and Secretariat. The more explicit steering should also include an explicit annual assessment of the performance of the Secretariat Director by the co-chairs and annual accounts approved by an external accountant.

***Recommendation 6*:** The impact of the Forum increases when results of Forum projects are applied broader than in the project setting itself. Although there is already quite some attention for result communication of research results (and this evaluation shows there is also quite uptake of these results in policy circles), there seems room to further boost the online visibility and uptake of both Forum publications and non-journal outputs. In case the Forum decides to increase capacity at the Secretariat, and a dedicated communications officer is appointed,

this could be one of the initial tasks. Developing "outcomes narratives" for each project and presenting them in annual reports, valorization reports, on a dedicated section of the Forum website, and/or on specialized platforms such as Kudos or Researchfish would increase exposure and consequently impact.



www.technopolis-group.com

technopolis group

September 2020

Belmont Forum Evaluation

Appendices

September 2020

Belmont Forum Evaluation

Appendices

Judith Vermeer (project leader), Etienne Vignola-Gagné, David Campbell, Ivette Oomens, Henrique Pinheiro, Liana Petrosova, Derek Eaton, Geert van der Veen

Table of Contents

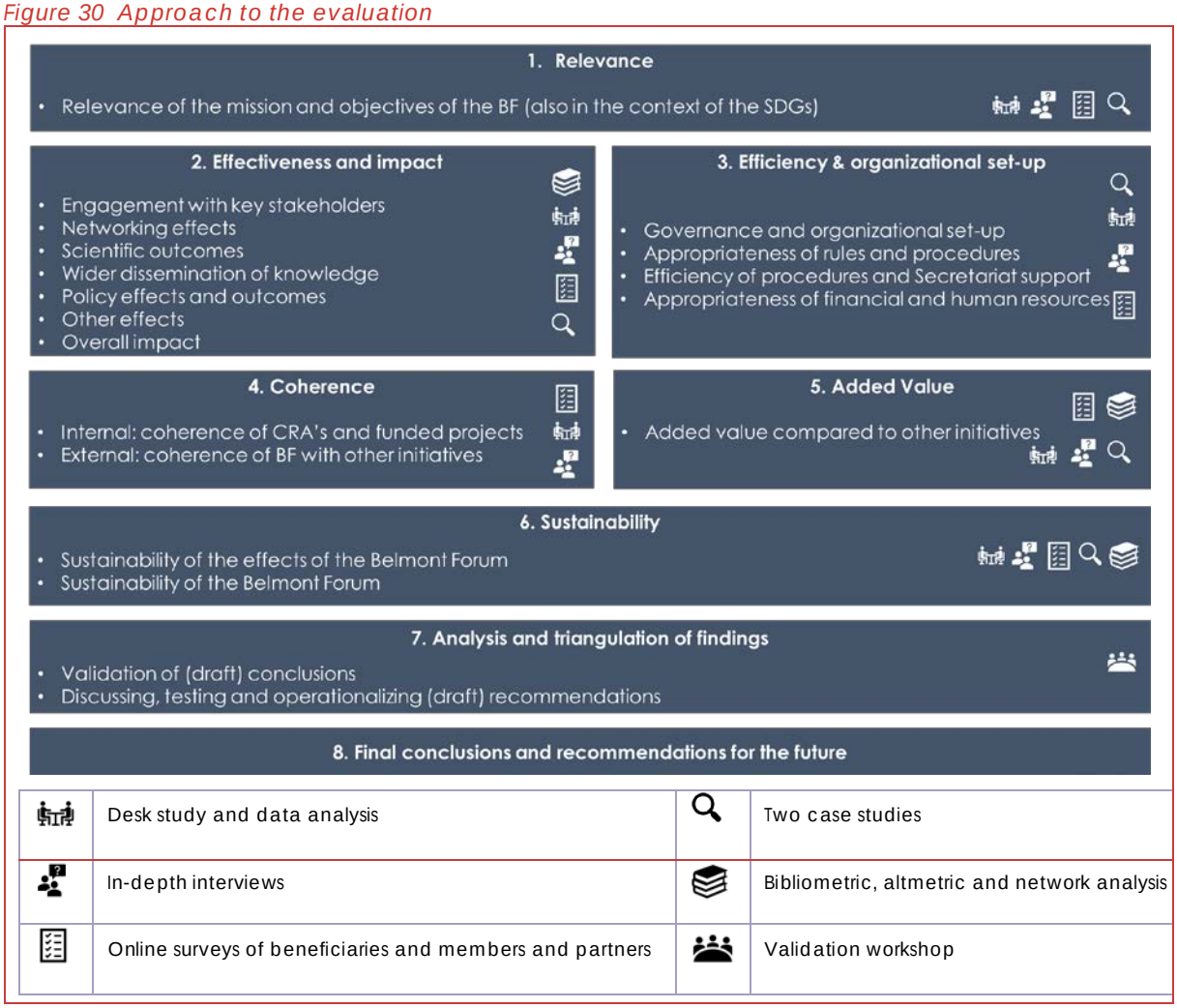
Survey analysis figures

Appendix A Methodological overview

A.1 Approach to the evaluation

The evaluation of the Belmont Forum (hereafter referred to as the Forum) consisted of a combination of quantitative and qualitative data collection and analysis tools. Figure 1 and table 1 summarise the main evaluation criteria and questions, and the methods used to answer these questions.

Figure 30 Approach to the evaluation



Source: Technopolis Group and Science-Metrix (2020)

Table 12 Overview of evaluation questions

#	Evaluation question
Relevance	
1	How relevant are the **mission and objectives** of the Forum (in the context of understanding, mitigating, and adapting to global environmental change, as well as achieving the SDGs)?
Effectiveness and impact	
2	How successful has the Forum been in effectively **engaging with its key (target) stakeholders**?

#	Evaluation question
3	**Networking effects:** How effective has the Forum been in generating new collaborations and partnerships across various sectors, disciplines, and countries around the globe? To what extent are the funded projects truly co-designed and co-created?
4	**Scientific outcomes:** To what extent did the Forum contribute to the science base for environmental change (understanding, mitigation, and adaptation)?
5	**Wider dissemination of knowledge:** How effective has the Forum been in disseminating knowledge and other outputs generated by the Forum? To what extent were results of the Forum disseminated, taken up and discussed beyond academic circles?
6	**Policy effects and outcomes:** To what extent did results of the Forum foster policy debate or developments at international and national level or facilitate policymaking / implementation?
7	**Other effects and outcomes:** Did the Forum have any other effects (intended or unintended, positive or negative)?
8	**Overall impact:** To what extent has the Forum (in collaboration with all its stakeholders) contributed to the Challenge of understanding, mitigating, and adapting to global environmental change?
Efficiency and organizational set-up	
9	How appropriate are the **governance and organizational set-up** of the Forum (especially considering its growing membership)?
10	Are the **rules and procedures** of the Forum appropriate (e.g. transparent and inclusive) and are they regularly reviewed for appropriateness?
11	To what extent has the Forum **carried out its work efficiently** (e.g. financial and human resources, internal procedures)?
12	Are the Forum's **financial and human resources** appropriate for the work entrusted to it?
Coherence	
13	**Internal coherence:** Are the CRAs and funded projects coherent with each other?
14	**External coherence:** Are the activities of the Forum coherent with other initiatives in the context of environmental change?
Added value	
15	What is the **added value** of the Forum (compared to other initiatives at various governance levels)?
Sustainability	
16	To what extent are the **effects and results** of (projects funded by) the Forum sustainable in the future?
17	To what extent is the Forum **sustainable in the future**?

Source: Technopolis Group and Science-Metrix (2020)

The tools and methods that we used to answer the evaluation questions can be summarised as follows:

- **Desk study and data analysis:** The evaluation team made use of and analysed existing relevant data and documentation (e.g. available through the Forum's website, BFgo, or provided to the evaluation team directly). The desk study included not only a review of relatively high-level strategic documentation [31] (on the mission, governance, and

[31] Such as the Forum's Memorandum of Understanding and Terms of Reference.

organization of the Forum), but also more operational documents (such as work plans, financial data, meeting minutes, and project data).

- **In-depth interviews:** The interview programme served to collect in-depth and detailed views and opinions from key stakeholders that are directly or indirectly involved in the Forum. In total, we conducted 40 interviews, namely 3 interviews with staff of the Secretariat, 20 members (incl. current and former members of the Steering Committee), 6 interviews with partner organizations, 5 interviews with external stakeholders, and 6 interviews with beneficiaries. The group of interviewed members also included people that were part of a TPO or GPC in the past. It should be noted that for some member organizations, we spoke with more than one representative. Appendix D contains the list of those interviewed and the detailed interview guides.
- **Online survey of beneficiaries:** The main purpose of the online survey of beneficiaries was to gain more insight into their views and opinions on topics such as the co-creation process of the Forum's funded projects, the proposal submission and granting processes, and the networking effects of the Forum. The survey was sent out to 1.600 beneficiaries via the Secretariat of the Forum on 11 March 2020. The deadline for the survey was 22 May 2020. The beneficiaries invited for the survey included both consortium leads and non-lead participants (i.e. partner PI's and other project participants). The Secretariat sent several calendar reminders to the beneficiaries in the weeks before the deadline. Additionally, survey participation was encouraged via messages on the Twitter and Linked-In accounts of the Forum. While sending the survey invitations, beneficiaries were made aware that the survey is anonymous. The survey consisted of 23 closed and 6 open questions. In total, we received 138 valid responses to the survey (8.6% response rate). Appendix B includes more detailed information on the profiles of the respondents (incl. country, type of organization, role in the funded project, etc.).
- **Online survey of members and partners:** While not originally planned under this evaluation, during the plenary meeting in Taipei we agreed with the members of the Forum to launch a short online survey among Forum members and partners. The main purpose of this survey was to obtain more quantitative data on certain views and ideas expressed by members and partners in the in-depth interviews. Thus, the survey helped to support the interview findings in a somewhat more substantiated manner. The survey was sent to the Forum's 29 member organizations and 6 partner organizations via the Forum Secretariat on 11 March 2020. In doing so, they were asked to provide one answer per member/partner organization. The Secretariat sent several calendar reminders to the members and partners in the weeks preceding the deadline. While sending the survey invitations and reminders, members and partners were made aware that the survey is anonymous, meaning that their responses could only be accessed by the evaluation team and would only be presented in aggregated ways. The deadline for the survey was 20 May 2020. The survey consisted of 23 closed and 9 open questions. In total, we received 18 valid responses to the survey, which represents 51% of all members and partners of the Forum. It should be noted that all members of the Forum have provided input to the evaluation, either via the in-depth interviews or the online survey. Appendix C includes more detailed information on the profiles of the respondents.
- **Case studies:** Two thematic case studies were undertaken to investigate the extent of, respectively, North-South collaboration and complementarity of the Forum with other relevant research funding initiatives. The main purpose of the case studies was to zoom in to these two key topics for the Forum and to gain a more in-depth understanding of the Forum's performance in terms of inclusiveness of the Global South and the Forum's collaboration and complementarity with other initiatives in the field. Both case studies relied

on a combination of information from the desk study, the survey of beneficiaries and in-depth interviews. For the case study on complementarity, the desk study compiled an overview of the landscape of relevant research collaboration and funding instruments. The survey and the interviews were formulated to include questions, where appropriate, covering one or both of the case studies. The case study on North-South collaboration also made use of results of the bibliometric analysis, which examined participation by global South in publications attributed to the Forum. Both cases studies are presented in Appendix E to this report.

- **Bibliometric, altmetric and social network analysis:** The Forum's vision is to support research consortia that contribute to traditional academic knowledge production (peer-reviewed publications), but also simultaneously foster societal outcomes within novel collaborations. The evaluation of project-level outcomes and outputs is therefore aimed to capture these two broad classes of outcomes, with the major constraint that information and data on the societal outcomes tend to be sparse, variable, and fragmented. By contrast, records on formal academic outputs are centrally and systematically codified within large global repositories. Evaluation of formal academic outputs enabled observations on research excellence achieved within supported projects (citation metrics), but also observations on a portion of the societal outcomes produced (including mentions in journalistic pieces, Wikipedia pages, or in regulatory science documents). Authorship of publications supported by the Forum also enabled tracking the degrees to which project teams have achieved inter- and transdisciplinarity, inter-sectoral collaborations, and North-South cooperation which are the key underlying (co-creation) dimensions of transdisciplinarity. When combined, findings from all of these measurements thus provide a rough assessment of transdisciplinary achievements, taking into consideration that there is currently no quantitative strategy available that single-handedly captures all component dimensions of transdisciplinarity (see section 3.4). The Forum publications were retrieved for the 2012–2014 CRAs (as scientific and societal outcomes of research require long periods to be fully realized). Benchmarking groups were assembled from:
    - prior publications (2007–2011) by the Forum awardees from the CRA 2012–2014;
    - Parallel, non-Forum publications issued by investigators supported by the Forum (also providing a counterfactual group when combined with findings on prior publications in a difference-in-differences analysis[32]);
    - publications supported by the BiodivERsA 2008 programme, a European-Union (EU) COST action with similar collaborative and societal engagement goals as the Forum;
    - major (i.e. associated with the highest volumes of publication) funders of research in the thematic areas where the Forum teams are active, including the NSF, UK Natural Environment Research Council (NERC), German Federal Ministry of Education and Research (BMBF) or European Research Council (ERC), among others[33];

[32] Buenstorf, G., & Koenig, J. (2020). Interrelated funding streams in a multi-funder university system: Evidence from the German Exzellenzinitiative. *Research Policy*, *49*(3), p. 103924. doi:10.1016/j.respol.2020.103924; European Commission. (2016). Counterfactual impact evaluation | EU Science Hub. *EU Science Hub*. Retrieved from https://ec.europa.eu/jrc/en/research-topic/counterfactual-impact-evaluation.

[33] In relation to the Global South, it should be noted that we computed most of our findings for the top 20 funders worldwide in the thematic space occupied by BF publications, including the Conselho Nacional de Desenvolvimento Científico e Tecnológico, National Natural Science Foundation of China, Chinese Academy of Sciences, Ministry of

- Bootstrapping was used to test for statistical significance where the Forum leads in performances relative to benchmarking groups were recorded.

Societal outcomes not associated with traditional research publications were recorded from the Belmont Forum Grants Operations (BFgo) project self-reporting data and project websites. For those outcomes available or described online, web citation statistics based on volumes of hyperlinks towards the resource were measured.

In addition to BFgo, this component of the evaluation (i.e. the bibliometric, altmetric and social network analysis) made use of the Scopus, PlumX, Overton and Uber suggest databases. Appendix E contains the full analysis.

- **Validation workshop:** The findings from the above-described tools and methods have led to preliminary conclusions to each of the evaluation questions. To check the validity of the findings and conclusions, validation workshop with members of the Forum Steering Committee took place on 30 July 2020. The validation workshop helped the evaluation team to further refine the final evaluation report and to draft recommendations that are useful and realistic for the Forum going forward.

A.2 Current challenges in evaluating transdisciplinarity and inter-governmental funding

Transdisciplinarity (TD) has been variously defined since the notion has started to emerge. TD projects can be characterized (drawing on both the Belmont Challenge but also broader scientific literature on TD)[34] by dimensions such as:

- Collaborative research with academic participants originating in multiple, diverse disciplinary contexts;
- Practising a co-constructive, systemic approach to research conduct where both (1) academic and (2) citizen, civic, policy or even business participants are engaged on an equal footing in defining work priorities for the project and the preferred pathways and instruments to reach them;
- Increasing collaboration across geographical areas and extending the co-design philosophy to relationships between geographical areas;
- Producing project outcomes that are directly useful, understandable and actionable by the citizen, civic, policy or even business participants and audiences, as much or even more so that traditional academic outputs (with the Belmont Challenge itself, for instance, placing research as "part of a value chain that is socially and ethically responsible and that fully involves all societal actors in the co-construction towards innovative solutions"

Conceptual frameworks for the evaluation of TD practices and programs are an object of research and innovation in their own right.[35] The formal program evaluation sector has yet to be able to integrate many of the proposals and practices put forward by academic research

---

Science and Technology of China, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Not all of these were presented in the final report (they are included in the introductory table of this technical report) because performances were not very different from those of other funders.

[34] Belmont Forum. (2016). *The Belmont Challenge: a global, environmental research mission for sustainability*. Doha; Schneider, F., Buser, T., Keller, R., Tribaldos, T., & Rist, S. (2019). Research funding programmes aiming for societal transformations: Ten key stages. *Science and Public Policy*, *46*(3), pp. 463–478. doi:10.1093/scipol/scy074; Belcher, B. M., Rasmussen, K. E., Kemshaw, M. R., & Zornes, D. A. (2016). Defining and assessing research quality in a transdisciplinary context. *Research Evaluation*, *25*(1), pp. 1–17. doi:10.1093/reseval/rvv025.

[35] B. M. Belcher et al., Defining and Assessing Research Quality in a Transdisciplinary Context.

on TD evaluation, however; it is also clear that evaluation contexts vary enormously from one to the next and that consensus cannot yet be achieved on a single way forward. Almost all of those proposals consulted by the evaluation team relied on expansive qualitative data collection strategies. [36] Additionally, the development of evaluation frameworks for TD programs is hampered by a lack of empirical data on prior examples and day-to-day management of these initiatives: "to date, very little documented experience exists in implementing TD research at the program level, and very few scientific studies have examined such programs beyond a focus on individual activities".[37]

In such a context, Technopolis has relied mainly on interviews, surveys and case studies that aim to capture the aspect of transdisciplinarity. More specifically, during the in-depth interviews as well as in the online surveys, the Forum's members, partners, and beneficiaries were explicitly asked about the nature of their activities, the stakeholders involved, and the level of transdisciplinary of funded projects.

Science-Metrix's expertise is very much focused on bibliometrics assessments conducted as part of program evaluation. The bibliometrics community broadly has yet to develop a single quantitative indicator suitable to track the complex/diverse interactions characterizing TD research (sets of indicators must be used for that purpose), or of robust tools for most types of possible societal outcomes from research. To meet the challenge of evaluating outputs and outcomes generated by the Forum-funded projects, Science-Metrix has relied on a combination of more established indicators and exploratory approaches. The use of a combination or dashboard of indicators was made necessary by the multi-dimensional character of TD projects. No single quantitative indicator can be expected to offer a comprehensive measurement of TD realization in research and knowledge transfer activities at this stage in the development of the field. In the end, the following indicators were included in the evaluation design to each provide insight into a discrete component of transdisciplinary project work:

- Interdisciplinarity index of peer-reviewed publications (measuring the integration of disciplinary diversity in new research finding)
- Multidisciplinarity index of peer-reviewed publications (measuring collaborative work that crosses disciplinary boundaries)
- shares (i.e. proportion expressed in percentage) of international co-publications within peer-reviewed publications and shares of North-South co-publications (measuring collaborative international work, particularly that brings together Global South and Global North countries)
- shares of inter-sectoral co-publications within peer-reviewed publications (collaborative work with research users and stakeholders)
- policy and science advice citations towards peer-reviewed publications (measuring work with a strong orientation towards the production of societal outcomes)
- shares of OA publications within peer-review publications; altmetrics profiles of peer-reviewed publications; web citations to videos, blog posts and other online content produced by project teams (measuring the degree of concern for broad diffusion of research findings and capacity to provoke curiosity in online communities)

---

[36] Belcher, B., Suryadarma, D., & Halimanjaya, A. (2017). Evaluating policy-relevant research: lessons from a series of theory-based outcomes assessments. *Palgrave Communications*, *3*(1), pp. 1–16. doi:10.1057/palcomms.2017.17.

[37] Schneider et al., Research Funding Programmes Aiming for Societal Transformations: Ten Key Stages.

These purely bibliometric indicators could also be combined with beneficiary survey findings and self-reported outcomes from awardee reports. In the end, a good amount of interpretative inference is necessary to combine multiple lines of evidence on TD achievements, with different lines associated with heterogeneous sources of data (bibliometric statistics vs descriptions of non-academic outcomes, for instance). This calls for a careful interpretation of study findings which is very much in line with the proposals found in other TD research evaluation schemes.[38]

It must be noted that the evaluation team has been able to retrieve incomplete evidence on societal outcomes of projects supported by the Forum. Outcome descriptions provided by project teams in their mid-term or final reports, although they hint at promising impacts, are often incomplete and unclear. Science-Metrix conducted additional data treatment of the BFgo database records on outcomes and outputs, and also documentary analysis on project websites and related sources to help better characterize these outcomes, often to little avail.

However, it must also be considered that evaluators with experience in tracking the societal outcomes of research often consider that these types of impact and outcomes may take as much as 10 or even 20 years to be fully realized,[39] which is beyond the scope of our analysis for the considered CRAs. Therefore, it is too early in this evaluation to obtain a comprehensive capture of the most far-reaching societal outcomes that may originate from the funded projects. This observation also provokes an interrogation as to whether project teams themselves are even able to appropriately witness and record the outcomes resulting from their research, greatly complicating attempts at monitoring.[40]

Research teams need support in better conceptualizing and describing the outcomes achieved, acknowledging that such outcomes can take several years to materialize. Future evaluations could focus entirely on tracking policy, cultural, social and economic outcomes from a selection of projects, given the magnitude of resources required and the complexity of such an endeavour.[41] This would most likely entail project-tailored interviews/surveys for projects completed at least 10 years prior.

A.3 Validity and limitations

The analysis using the tools and methods as described in the previous section were mostly carried out as planned. Nonetheless, as most evaluation studies do, this study also faced some limitations and challenges related to data availability and stakeholder response rates and biases. Where relevant, these limitations are pointed out in the description of the evaluation findings. This section describes the main general limitations of this evaluation study.

First, a (small) limitation relates to the **number of interviews conducted**. While we had foreseen to conduct a total of 45 interviews, we were only able to conduct 40 interviews. Despite several follow-ups by email and telephone, it seemed particularly difficult to engage certain members of the Forum, as well as beneficiaries. A potential explanation for the lack of responsiveness may be the COVID-19 crisis (and shifting priorities resulting from this crisis). We attempted to mitigate this limitation as much as possible, among others by involving the Secretariat of the

[38] B. M. Belcher et al., Defining and Assessing Research Quality in a Transdisciplinary Context.

[39] Langfeldt, L., & Scordato, L. (2015). *Assessing the broader impacts of research. A review of methods and practices.* Oslo: Nordic Institute for Studies in Innovation, Research and Education. Retrieved from https://nifu.brage.unit.no/nifu-xmlui/handle/11250/282742.

[40] Schneider et al., Research Funding Programmes Aiming for Societal Transformations: Ten Key Stages.

[41] B. Belcher et al., Evaluating Policy-Relevant Research: Lessons from a Series of Theory-Based Outcomes Assessments.

Forum in boosting responses to the interviews. Given the additional survey of members and partners that was introduced at the inception phase of this study, we feel that sufficient input was received from the Forum's members. In the end, we were able to ensure that all members provided input to the evaluation, either via an interview or via a survey.

Second, the **number of responses of beneficiaries to the survey** were somewhat lower than we had expected, given the relatively long period available to respond to the survey (more than 10 weeks). Related to this, there were also some biases in the responses to the online survey of beneficiaries. The survey generated a relatively low number of responses from countries in the Global South, and from beneficiaries of early CRAs. While we aimed to boost response rates of these specific groups of beneficiaries, the results are still suboptimal. To mitigate the effects of this (especially in our assessment of the level of inclusiveness of the Forum and the participation of the Global South), the survey findings were complemented by the in-depth interviews and a specific case study on this topic.[42]

Third, the bibliometric, altmetric, and network analyses faced challenges concerning their scope with respect to societal outcomes of the research. The Forum's mission puts a clear emphasis on support for **transdisciplinary research**. As already mentioned above, bibliometric and other indicators used here provided quantitative insights into multiple, discrete components of transdisciplinary practice, but cannot offer a single indicator of transdisciplinarity achievement.

A partial assessment of the Forum's achievements in transdisciplinarity can be obtained by juxtaposing these discrete measurements. It can also be noted that the discrete components of the transdisciplinarity concept dealing with intellectual diversity (interdisciplinarity and multidisciplinary) were regrouped under the term of cross-disciplinarity when discussed generically. A definitive evaluation of the extent to which co-design and co-implementation were achieved in supported projects cannot be fully realized purely with the quantitative strategies currently available in bibliometrics. However, findings originating from cross-disciplinary, international and inter-sectoral co-publications, as well as from qualitative observations on societal outcomes and some answers to the beneficiary survey provided partial insights into this dimension.

Concerning the **evaluation of societal outcomes of research projects**, one would most often use extensive *qualitative* methodological approaches, including case studies, interviews, and expert panels. While this evaluation did include survey, interview and case study components, a comprehensive evaluation of societally oriented outcomes would again necessitate a dedicated study. Moreover, such evaluation would need to take place at least some years after the end of projects, to allow for societal effects to materialise (taking as many as 10 or more years).[43] To offer *quantitative* findings on potential early societal uptake of the Forum project outcomes, we notably collected web citation, hyperlink and usage statistics on non-journal outputs supported by the Forum. Outputs such as online videos, blog posts, web pages, webinars, policy reports and journalistic pieces were expected to receive some attention from society-side users and stakeholders, the attention that can be captured in online links, downloads and/or views. This approach must be considered experimental and has yet to be systematically used in formal evaluation contexts. With the URLs of citing websites available, it was hoped that citations of special interest, originating from local communities, non-

[42] It should be noted that in the survey analysis, we explored whether there were any differences in the answers of beneficiaries from the Global North and beneficiaries from the Global South. Only in very few instances, such differences were found. These are reported separate graphs in Appendix B.

[43] Langfeldt, & Scordato, *Assessing the broader impacts of research. A review of methods and practices*.

governmental organisations (NGOs), partners, and the like, could be identified. This portion of the analysis retrieved comparatively low numbers of observations, and for the most part, did not help identify cases of in-depth participatory or local uptake. It is unclear if altmetrics methods are not yet appropriate for use with non-journal outputs, or whether this is indicative of restricted outcomes here. To reduce the uncertainty surrounding this issue, the findings from the altmetric analysis were complemented and triangulated with the other methods used for this evaluation, in particular the online surveys and stakeholder interviews.

It should be noted that findings for **benchmarking groups** provided here can be used to assess Forum performances only. Comparison between benchmarking groups can in no way be considered valid. Tables presenting bibliometric findings can in no way be read as rankings including performances from multiple funders. This is particularly true given that statistical tests on the significance of differences in performances were only performed for differences in measurements between the Forum and benchmarking groups, and never between benchmarking groups.

In research evaluation – where bibliometric assessments typically rely on random samples of large cohorts of researchers, or on the entirety of a program's cohorts as in this study – the use of formal statistical testing is also warranted to enable statistical inference on future cohorts of the corresponding programs, assuming the available data is representative of these latter cohorts (implying that program characteristics and implementation remain mostly stable). Where relevant, hypothesis testing was performed to assess the statistical significance of the observed differences and differences-in-differences. Robust p-values were estimated using a **bootstrapping procedure** (unless otherwise stated) instead of traditional statistical testing.[44] This procedure successfully addresses some of the challenges of working with bibliometric data, in a better way than would be feasible with traditional statistical testing (see section the technical appendix for full details). Apart from the p-values, the bootstrapping procedure was also used to estimate 95% confidence intervals of the bibliometric indicators as well as of differences (or differences-in-differences) between groups. These intervals are sometimes shown in brackets below point estimates in this study's tables. Note that due to space limitations, the differences between groups are not reported in this study's tables as they can be inferred from the scores of the groups being compared. P-values are reported for the difference between the average score for the Belmont set of publication and each benchmarking group. In this study, low p-values suggest that the Forum papers scores are, on average, higher than the relevant benchmarking group."

The **p-values** reported in this study's tables were highlighted with symbols whenever they are smaller than the customary threshold of significance (*: $p < 0.01$; †: $p < 0.05$; ‡: < 0.1). However, the p-values and other statistics should be interpreted while keeping in mind that this report is part of a much larger investigation in which multiple lines of evidence are used. Accordingly, if a p-value is larger than 0.05 but smaller than 0.15, we would still argue that the observed difference is likely to be observed in future cohorts of the program, assuming that their characteristics remain roughly unchanged and that the program does not drastically change. In that case, applying the principle of the convergence of partial indicators through triangulation with the evaluation's other lines of evidence helped derive robust conclusions.

Finally, the databases used for bibliometric, altmetric and social network analyses all tend to be **biassed towards English-language publications**. The PlumX, Overton and Ubersuggest databases are emerging sources that have yet to be fully characterised for potential biases in

[44] Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. *An Introduction to the Bootstrap*. doi:10.1007/978-1-4899-4541-9.

global representativity or recall of their records. Annual project reports in BFgo, which notably integrate records from self-reported project outcomes and outputs, display variability in information reliability and completeness on a project-by-project basis. BFgo records were enriched by online queries and parsing of project websites, but not all Forum-funded projects had created or maintained a project website.

Regarding the **overall validity and representativeness of this evaluation study**, we feel that despite the above-described challenges and limitations, we were able to draw reliable and evidence-based conclusions and useful recommendations for the future. The evaluation draws on several different sources and has consulted a wide range of stakeholders that are either directly or indirectly involved in the Forum. Moreover, the triangulation of data and information from the various sources allowed us to enhance the reliability of the evaluation findings and to identify and eliminate any erroneous findings coming from single tools or methods.

Appendix B Survey of the Beneficiaries of the Belmont Forum

The beneficiaries of the Forum were surveyed in March-May 2020. The online survey was distributed to beneficiaries via an open link and was open for responses for 11 weeks. Survey invitations and reminders were sent via the Secretariat of the Forum. The survey was sent to all named project participants of all CRAs as of March 2020, which accounts for more than 1500 persons. Email addresses from project proposals and reports were used. Only about 150 of the email addresses were invalid or inactive; so, the message was received by approximately 1400 project participants.

Responses to the survey were fully anonymised. However, participants were allowed to leave their contact details if they were open to further communication with the study team. From those who left their contact details, 14 respondents were approached to conduct further in-depth interviews. They were selected based on several criteria including:

- Respondents with a relatively smaller amount of "I don't know" answer category answers. This ensured that respondents had meaningful contributions to the study questions.
- Respondents from different geographies
- Respondents with various experience and attitudes towards the Forum

In total, we have received 237 responses to the survey. After removing insufficiently filled partial responses (87 filled less than half of the questionnaire) and responses from those who did not receive funding (12), 138 remaining responses were analysed using Microsoft Excel. Approximately 16% of those contacted have responded to the survey. The filtering and removal of responses were done in close consultation with the Belmont Forum. The low response rate is can be explained by several issues:

- The surveys were sent around amid COVID-19 outbreak which meant that for many of the respondents' other issues have become more of a priority
- Many recipients have been involved with the Forum some time ago and do not have the connectivity with Belmont Forum
- Survey conflicted with the Forum's annual reporting requirement.

In the remainder of this appendix, we first provide an overview of the survey questions followed by an overview of the results of closed questions with annotated graphs.

A.4 Survey questions

Topic	#	Survey Question	Answer options
Section A: Respondent characteristics	A1	In which country are you employed?	Select from the list of world countries
	A2	What type of organisation do you work for?	Select from: university, research institute, NGO, company, other
	A3	What is your status in the consortium?	Select from: coordinator (or lead PI), partner (or non-lead investigator)
	A4	For which CRAs (Collaborative Research Actions) did you receive funding? • List of CRA provided	Select from: I received funding for this CRA; I applied but did not eceive funding for this CRA

Section B: Rationale and mission of the Belmont Forum .	B1	To what extent do you feel the activities of the Belmont Forum contribute to understanding, mitigating, and adapting to global environmental change?	Select from: 1 - not at all; 2 - to a small extent; 3 - to some extent; 4 – to a large extent; 5 - to a very large extent; Not applicable; Don't know
	B2	How would you rate the level of alignment between: • The problems and needs to adapt to global environmental change and the mission of the Belmont Forum • The problems and needs to adapt to global environmental change and the CRAs of the Belmont Forum • The problems and needs to adapt to global environmental change and the selected projects by the Belmont Forum	Select from: 1 – not aligned at all; 2 – aligned to a small extent; 3 – aligned to a moderate extent; 4 – well aligned; 5 – very well aligned; Don't know
Section C: Organisational aspects of the Belmont Forum	C1	How would you rate the Belmont Forum in terms of: • Quality of the application site • Professionalism of the Belmont Forum	Select from: 1- very poor; 2 – poor; 3 – adequate; 4 – good; 5 – very good; Don't know
	C2	How satisfied are you with the: • Frequency of interaction of the Belmont Forum • Administrative burden related to answering to the needs of the Belmont Forum • Quality of the interaction with the Belmont Forum	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
Section D: CRAs (Collaborative Research Actions)	D1	To what extent do you feel that the time spent on a proposal for the Belmont Forum is proportional to the success rate?	Select from: 1 - not at all; 2 - to a small extent; 3 - to some extent; 4 – to a large extent; 5 - to a very large extent; Not applicable; Don't know
	D2	How satisfied are you with the efficiency of: • The scoping process of CRAs • The submission process of proposals • The evaluation of proposals • The selection of proposals • The overall award process	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D3	How satisfied are you with the level of inclusiveness of the Belmont?	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D4	How satisfied are you with the level of transparency of: • The scoping process of CRAs • The application process • The evaluation of proposals • The selection of proposals • The award process	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D5	How satisfied are you with the: • Monitoring process of Belmont Forum projects • The mid-term review • The final review of Belmont Forum projects	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know

	D6	Do you have any suggestions for improvement of Belmont Forum procedures?	Open answer box
Section E: Effectiveness and impact	E1	To what extent was your project transdisciplinary (i.e. co-developed and co-implemented by researchersby researchers from natural sciences, social sciences and humanities, and with the involvement of stakeholders)? What was the added value of this transdisciplinary? Please elaborate.	Open answer box
	E2	How likely is it that your current research collaboration would have started without funding of the Belmont Forum?	Select from: 1 - very not likely; 2 - not likely; 3 - somewhat likely; 4 – likely; 5 - very likely; Don't know
	E3	How would you rate the likelihood of: • Continuation of your project after the Belmont Forum funding ends • Continued collaboration with your consortium partners (in the same or new projects) • Use of research results of your project by yourself after the projects ends • Use of research results of your project by others than yourself or your project partners after the projects ends	Select from: 1 - very not likely; 2 - not likely; 3 - somewhat likely; 4 – likely; 5 - very likely; Don't know
	E4	To what extent has your project involved co-design or co-creation with researchers from the Global South?	Select from: 1 - not at all; 2 - to a small extent; 3 - to some extent; 4 – to a large extent; 5 - to a very large extent; Not applicable; Don't know
	E5	If the Belmont Forum did not exist, how likely would it be that....? • You would have had access to research funds to support our project • You would have collaborated with the same (or similar types of) partners • You would have achieved the same (or similar types of) scientific results	Select from: 1 - very not likely; 2 - not likely; 3 - somewhat likely; 4 – likely; 5 - very likely; Don't know
	E6	Do you feel that your project has had a significant impact on societal challenges (such as people's well-being, climate, etc.)?	Select from: Yes; No; Don't know
	E7	Do you feel that your project has had a significant scientific impact?	Select from: Yes; No; Don't know
Section F: Communication and dissemination	F1	How would you rate the quality of the Belmont Forum communication tools? • Website • Twitter • Instagram • LinkedIn • Newsletters • YouTube channel • Belmont Forum Grant operations (GO) system • Guidance documents (e.g. conflict of interest, application guidance) • Meetings/events of Belmont Forum (if you attended any)	Select from: 1 – very low quality; 2 – low quality; 3 - moderate quality; 4 - good quality; 5 – high quality; Don't know

	F2	What, if any, events/meetings organised by the Belmont Forum have you attended?	Open answer box
	F3	What activities did you undertake to disseminate your project results?	Tick from: Via scientific publications; Via non-scientific publications; By giving a presentation at a conference/event; By presenting my results in a webinar; By presenting my results in discussions with policy makers; Other
	F4	Do you feel the Belmont Forum provides you with incentives to disseminate your project outputs?	Select from: Yes, the Belmont Forum provides formal incentives; Yes, the Belmont Forum provides informal incentives; Yes, the Belmont Forum provides both formal and informal incentives; No
	F5	How would you rate the usefulness of the following: • Data-sharing tools and infrastructure offered by the Belmont Forum • The open access policy and requirements of the Belmont Forum • The trainings provided by the Belmont Forum • The valorisation activities provided by the Belmont Forum	Select from: 1 – not useful at all; 2 – slightly useful; 3 - moderately useful; 4 - useful; 5 – very useful; Don't know
Section G: Conclusion	G1	Do you have any other comments regarding this survey?	Open answer box

A.5 Survey results

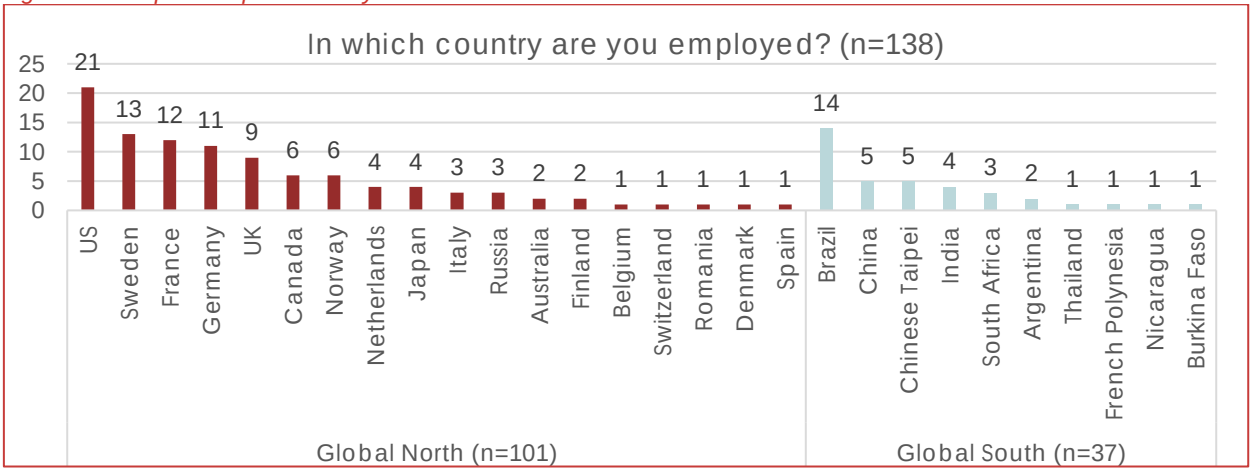
This section presents a basic analysis of the closed (multiple-choice) questions in the beneficiaries' survey. Responses to open questions are not presented. Questions are presented in the order as outlined in the table above. Brief descriptions are provided as annotations to the results presented in figures.

A.5.1 Section A: Respondent characteristics

Representatives of the Global North dominate the composition of respondents of the beneficiaries' survey. This finding is in line with the expectations and is proportionate to the overall number of the beneficiaries of the Belmont Forum.

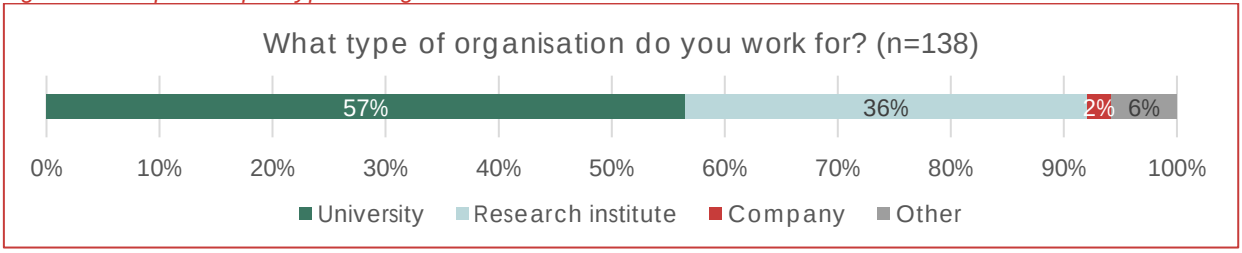
Further in this analysis, separation into North and South is only presented when deemed necessary. However, the study team has performed a segregated analysis for each of the survey questions. The differences in perception between respondents from North and South were insignificant, and overall followed the same trends for both groups. Also, the number of responses can be considered too low to make any valuable conclusions on the difference in perception between these two groups.

Figure 31 Responses per country



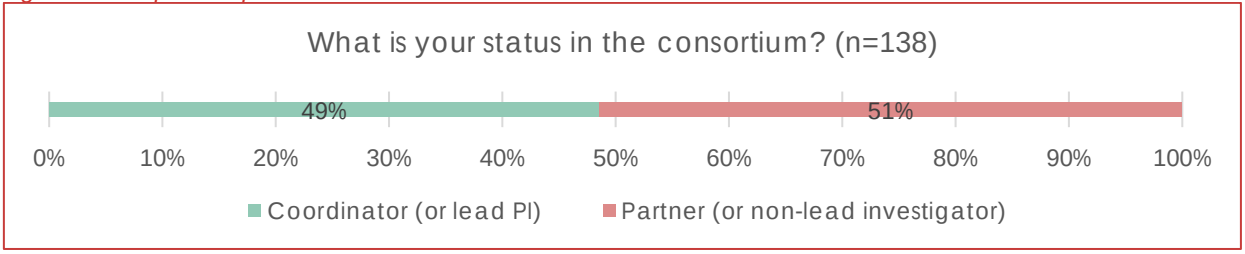
The vast majority of respondents are employed by universities and research institutes. The category "other" includes governmental bodies, non-governmental organisations (NGO), think tanks, and self-employed respondents.

Figure 32 Responses per type of organisation



There is a good balance in terms of respondents' role in their respective research consortia.

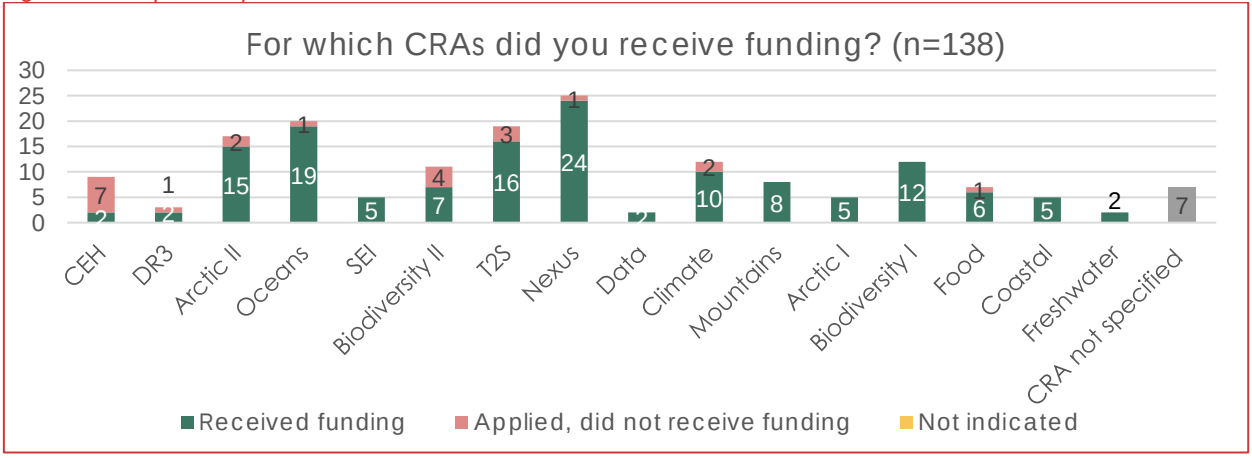
Figure 33 Responses per consortium role



The Figure below presents the survey participation per CRA. CRAs are presented in the order from most recent (left) to oldest (right). Overall, we received more responses from those who participated in the more recent CRAs, which is logical given that these responses are more likely to be presently engaged with the Forum and can share more relevant and meaningful insights. Most of the respondents received funding, while 7 respondents did not answer this question (did not indicate whether they received funding for any CRA). However, it was decided to include these respondents in the further analysis as it cannot be assumed that they did not receive any funding at all as they did not specify that.

It must also be noted several respondents (12) indicated that they have applied but have not received funding to any of the CRA. Those respondents were excluded from the analysis, which is pointed out in the introduction of Appendix B.

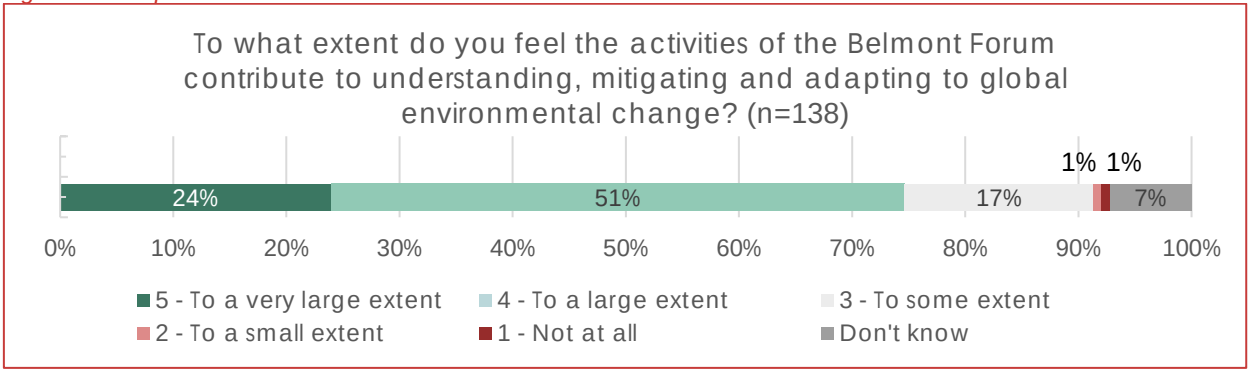
Figure 34 Responses per CRA



A.5.2 *Section B: Rationale and mission of the Belmont Forum*

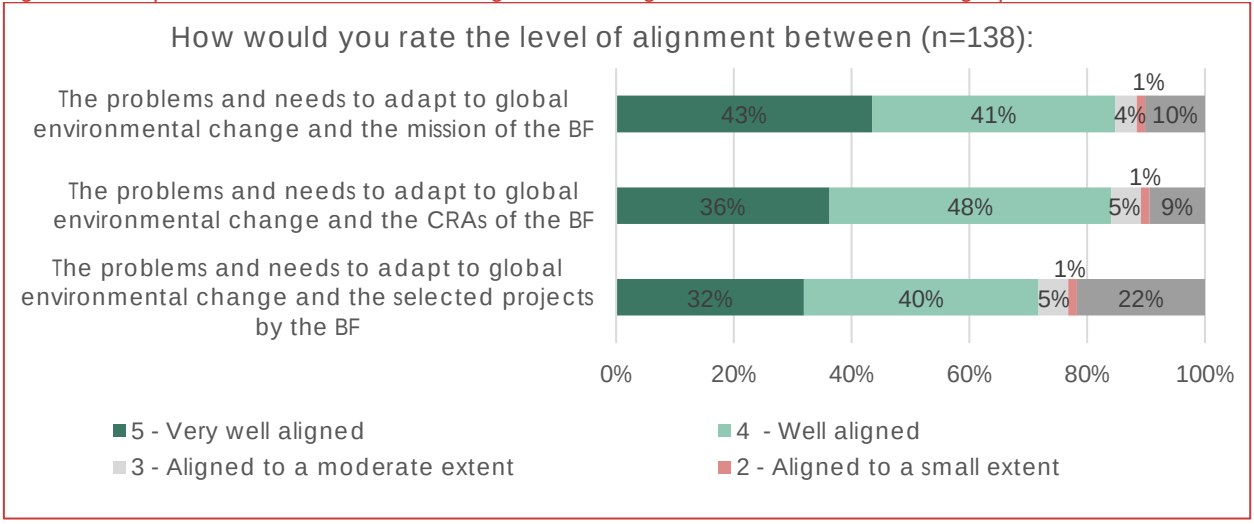
Overall respondents were positive about the contribution of the Forum's activities to understanding, mitigating, and adapting to global environmental change. A large majority (75 %) indicated that the Forum contributed to a large or very large extent to these goals.

Figure 35 Respondents' view on the relevance of the Forum



Similarly, the alignment between the Forum's mission and the global environmental change problems is rated highly (Figure 6). Respondents believe that both the CRAs and project funded under them are well aligned with the problems and needs to adapt to global environmental change.

Figure 36 Respondents' view on Forums alignment with global environmental change problems



A.5.3 *Section C: Organisational aspects of the Belmont Forum*

The majority of survey respondents rate the quality of the application site and professionalism of the Forum as good or very good (Figure 7). Similarly, about 80% of the respondents are generally satisfied or very satisfied with the frequency and quality of interaction with the Forum as well as with the administrative burden related to the needs of the Forum (Figure 8). It should be noted that the level of administrative burden is subjective to the role in the consortium: lead investigators often take up all the administrative tasks. Also, since the funding is provided through national organisations, additional administrative requirements of those organisations may apply. Here we assume that the participants differentiate the burden associated with the Forum versus that associated with national funding organisations.

Figure 37 Respondents views on the quality and professionalism of the Belmont Forum

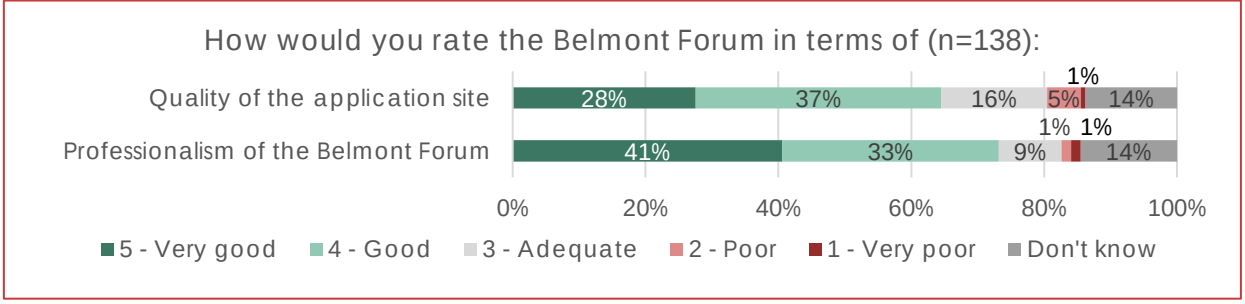
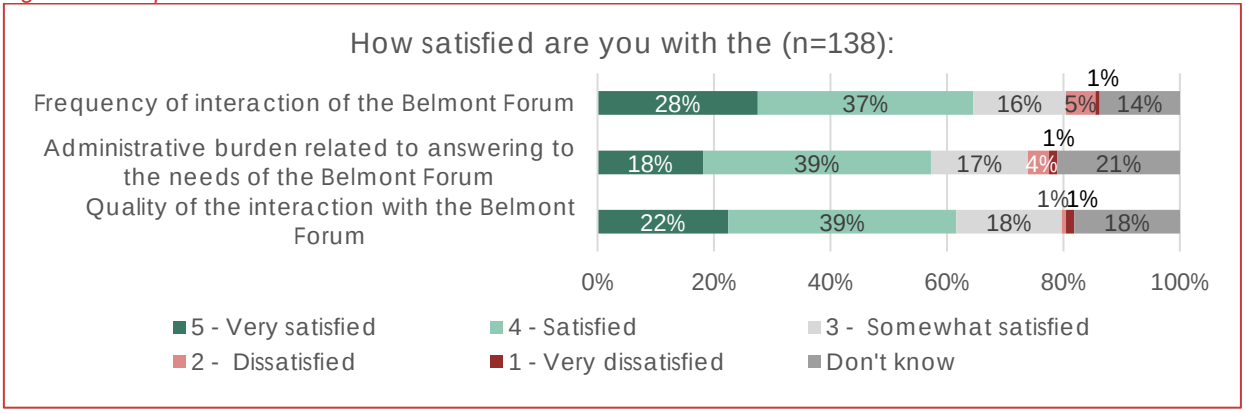


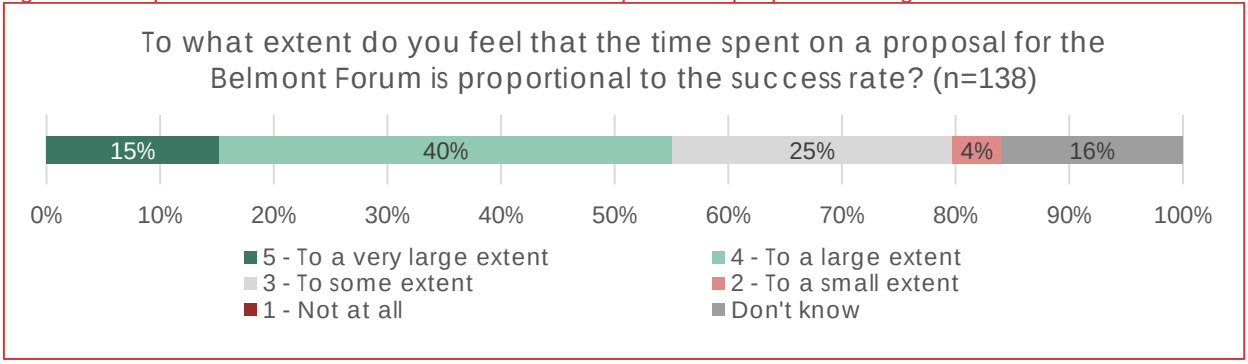
Figure 38 Respondents views on interactions with Belmont Forum



A.5.4 *Section D: CRAs (Collaborative Research Actions)*

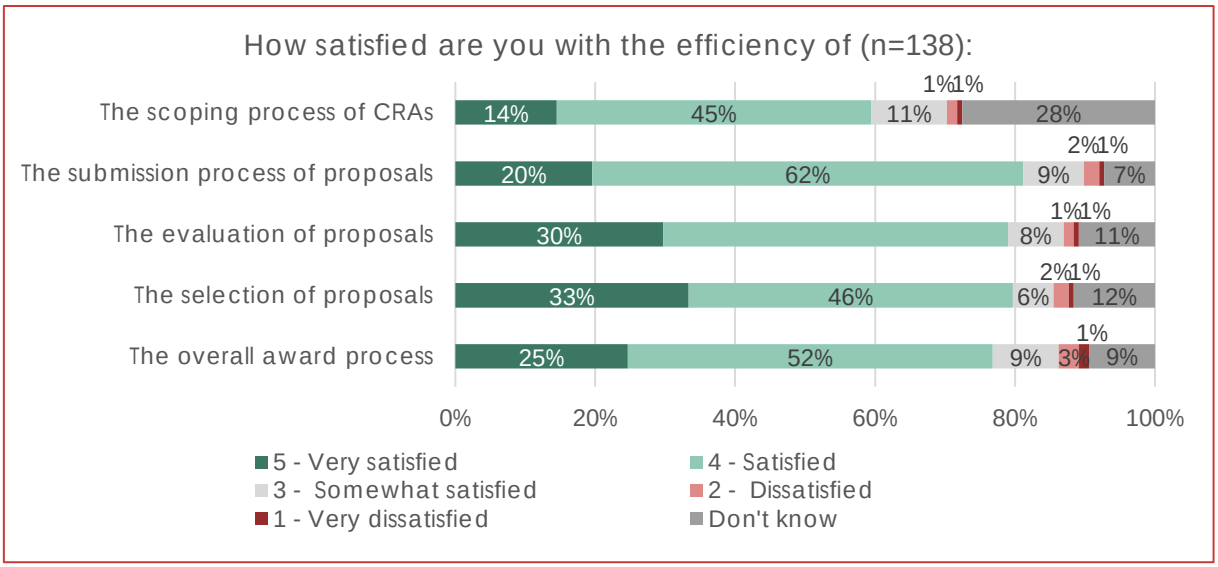
Over half of the respondents are satisfied with the ratio between the efforts needed for writing proposals for Forum and the success rate. However, the results may be skewed since the majority of respondents indicated that their funding applications have been successful, thus they are more likely to be more positive on this subject.

Figure 39 Respondents' views on the level of effort required for proposal writing



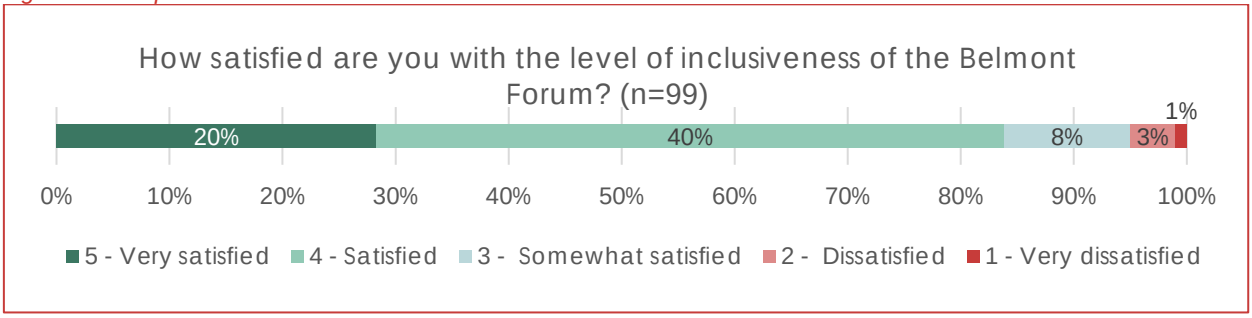
Overall respondents are in majority satisfied or very satisfied with the efficiency of the proposal-related processes. A somewhat lower number of respondents seem to be aware of the specifics of the CRA scoping process, given that 28% indicated "Don't know".

Figure 40 Respondents' views on the efficiency aspects of CRA development and proposal evaluation



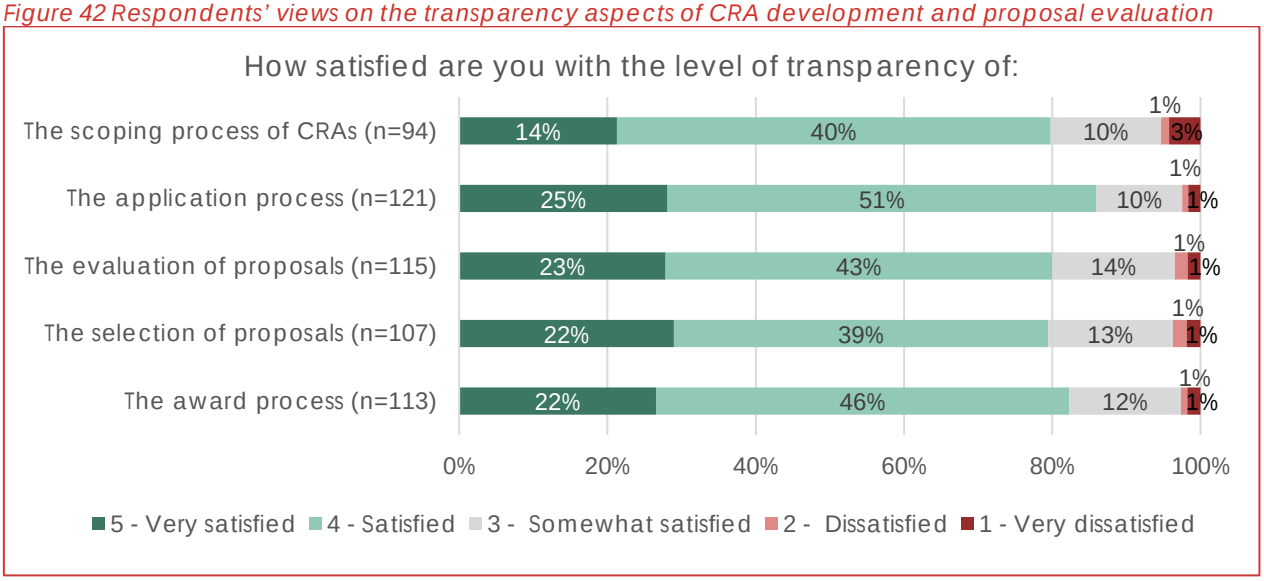
Most respondents are satisfied with the inclusiveness of the Forum. However, the analysis does not include a number of respondents (n=39) who indicated that they are not (sufficiently) aware of the Forums inclusiveness efforts.

Figure 41 Respondents' views on the level of inclusiveness of the Belmont Forum



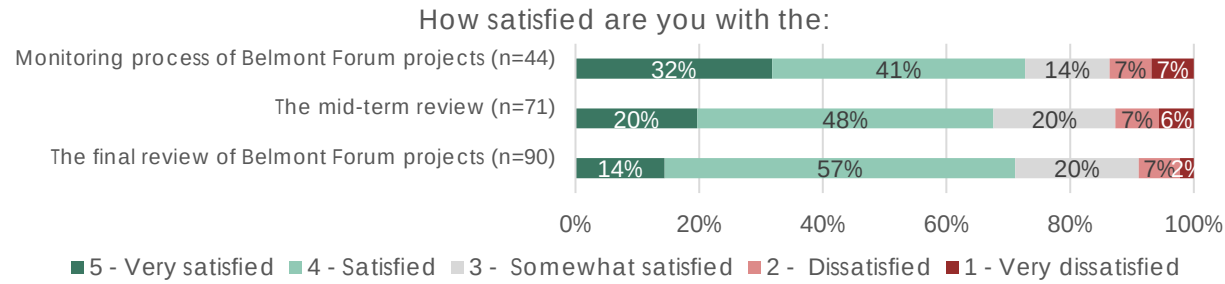
Results for satisfaction with transparency are similar to those presented in Figure 10: overall most respondents are satisfied or very satisfied with most aspects of transparency. However, here as well, respondents who selected "Don't know" option, were excluded from the analysis.

Figure 42 Respondents' views on the transparency aspects of CRA development and proposal evaluation



In terms of the Forum's monitoring and evaluation processes, only a small number of respondents could provide an answer, while the rest of the respondents selected "I don't know" option. As presented in Figure 13 (category "I don't know" is removed from the analysis", the majority of respondents are satisfied and/or highly satisfied with the monitoring and review of projects. The survey results also reveal that a large proportion of respondents are not sufficiently aware of monitoring and review processes. This may be because respondents are a) not lead project coordinator and are therefore not sufficiently involved in the monitoring and evaluation; b) respondents' projects are still at the early stages of implementation. Nonetheless, for those respondents who did rate the monitoring and evaluation, the satisfaction rates are exceptionally high.

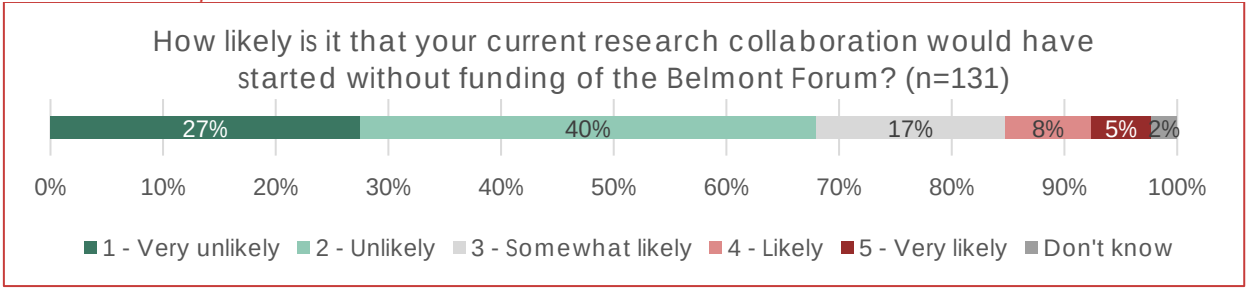
Figure 43 Respondents' views on the monitoring and evaluation of Belmont Forum's projects



A.5.5 *Section E: Effectiveness and impact*

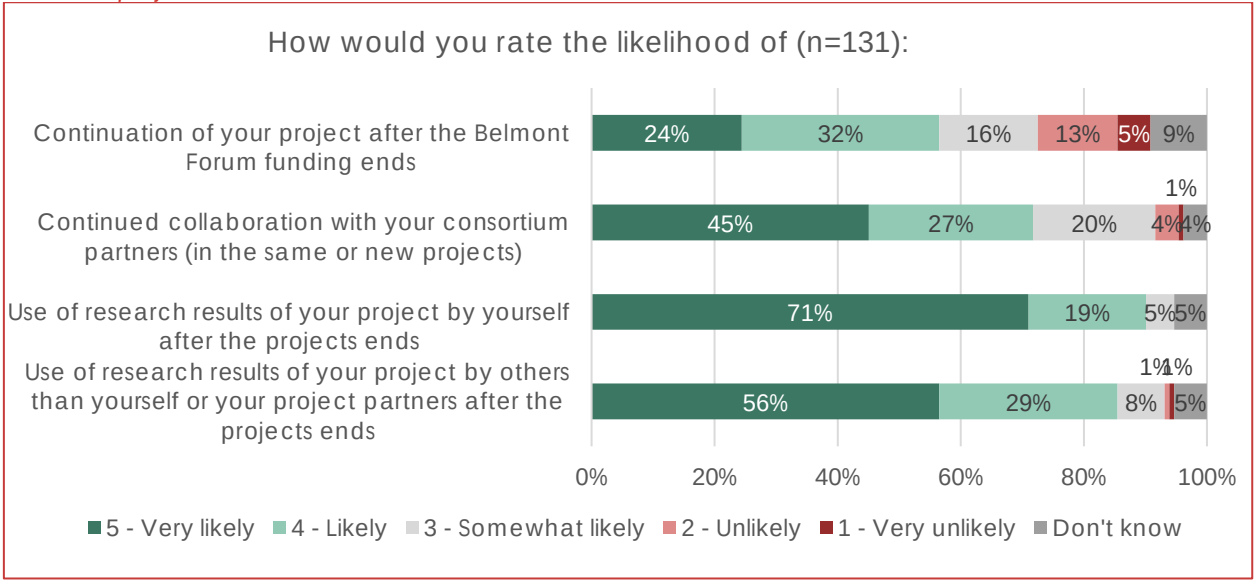
The majority of respondents believe that the Forum funding was a unique opportunity to get their collaborations started: almost 70% indicates that the current research collaboration was unlikely to have started without the funding of the Forum (Figure 14).

Figure 44 Respondents' views on the uniqueness of Belmont Forum's funding opportunity for the development of their research consortium



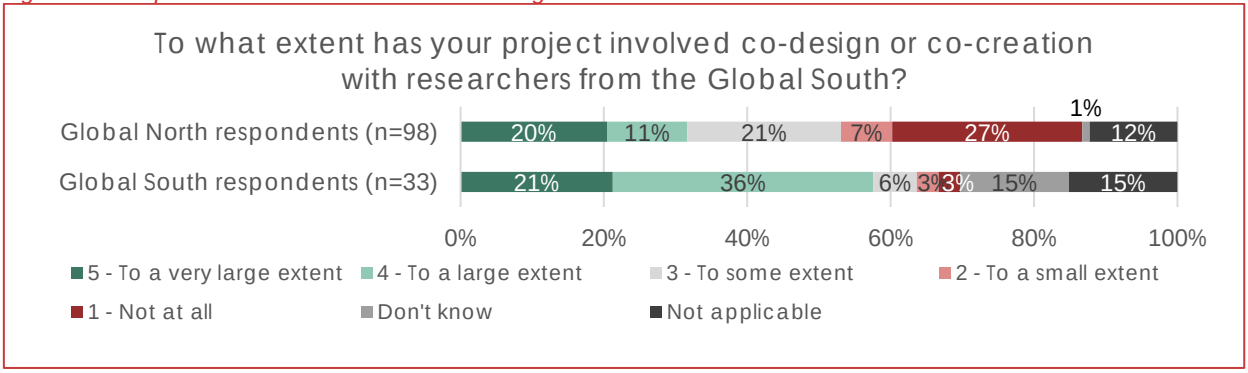
While slightly half of the respondents believe in the likelihood of project continuation beyond the Forum's funding, a much higher number of respondents believe in continued collaboration within the consortium and use of project results beyond it.

Figure 45 Respondents' views on the likelihood of continued collaboration beyond their respective projects



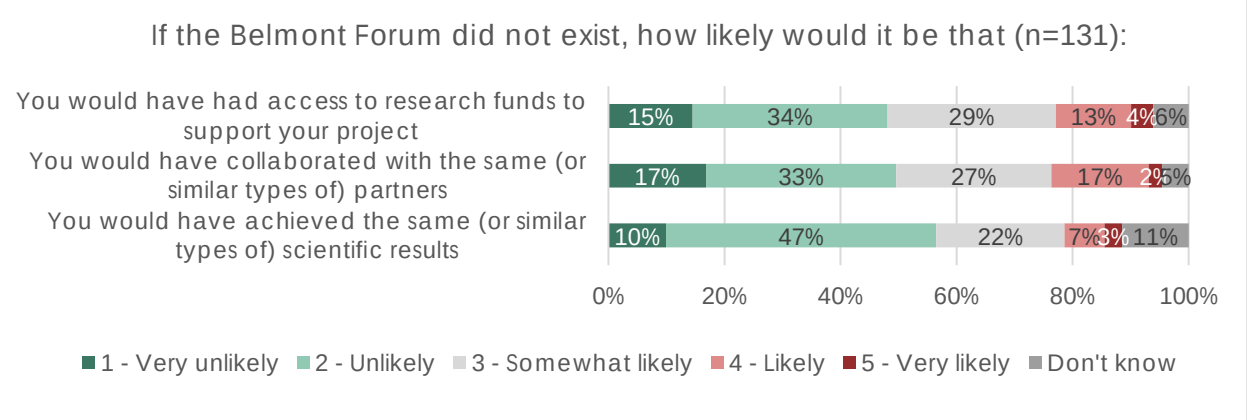
Among the survey respondents, the level of co-design and co-creation with researchers from the Global South was lower among Global North respondents, of which 31% indicated to a large extent or more such activities with researchers from the Global South. Respondents from the Global South did involve co-design and co-creation with the Global South more widely, with 57% of respondents indicating to a large extent or more.

Figure 46 Respondents' level of use of co-design and co-creation with researchers from Global South



Most of the respondents believe it would be unlikely to have access to the same resources, partners and achieve the same scientific results without the Forum's funding.

Figure 47 Respondents' views on the influence of Belmont Forum's funding on their project support and results.



In terms of scientific and societal impact, more than one-third of respondents could not specify whether their project could make a significant contribution. This is explained by the fact that for many respondents the projects are not yet at the stage where such impacts could be seen. In addition, an overall assessment of such impacts is challenging. Nonetheless, most respondents who did have insights into their project's impact on societal challenges and science did believe that their projects have had significant scientific and societal impacts.

Figure 48 Respondents' views on the impact of their projects on societal challenges

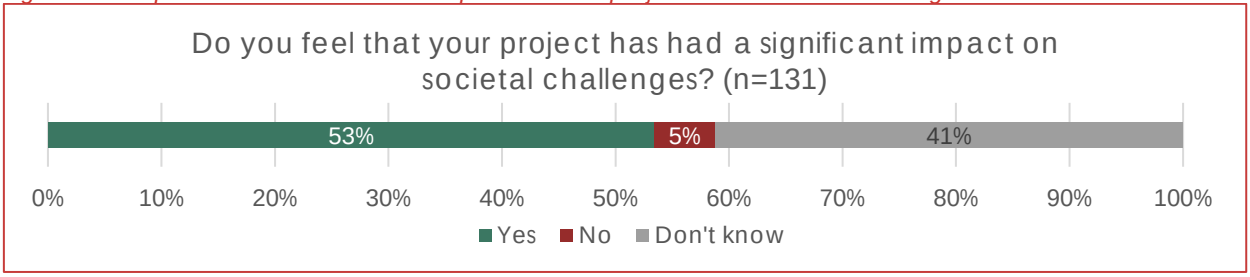
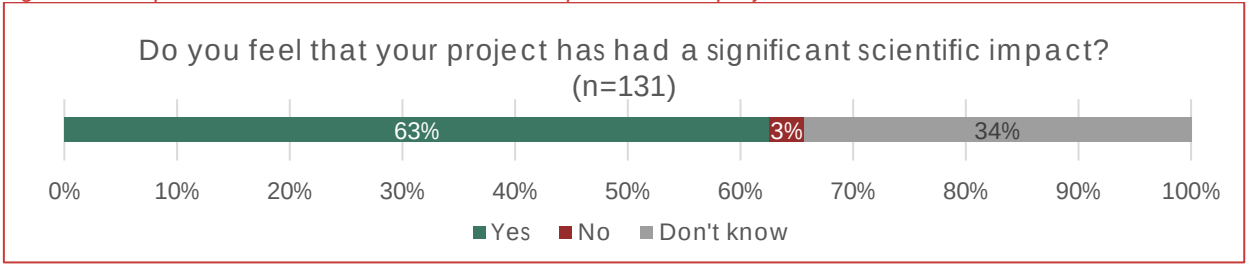


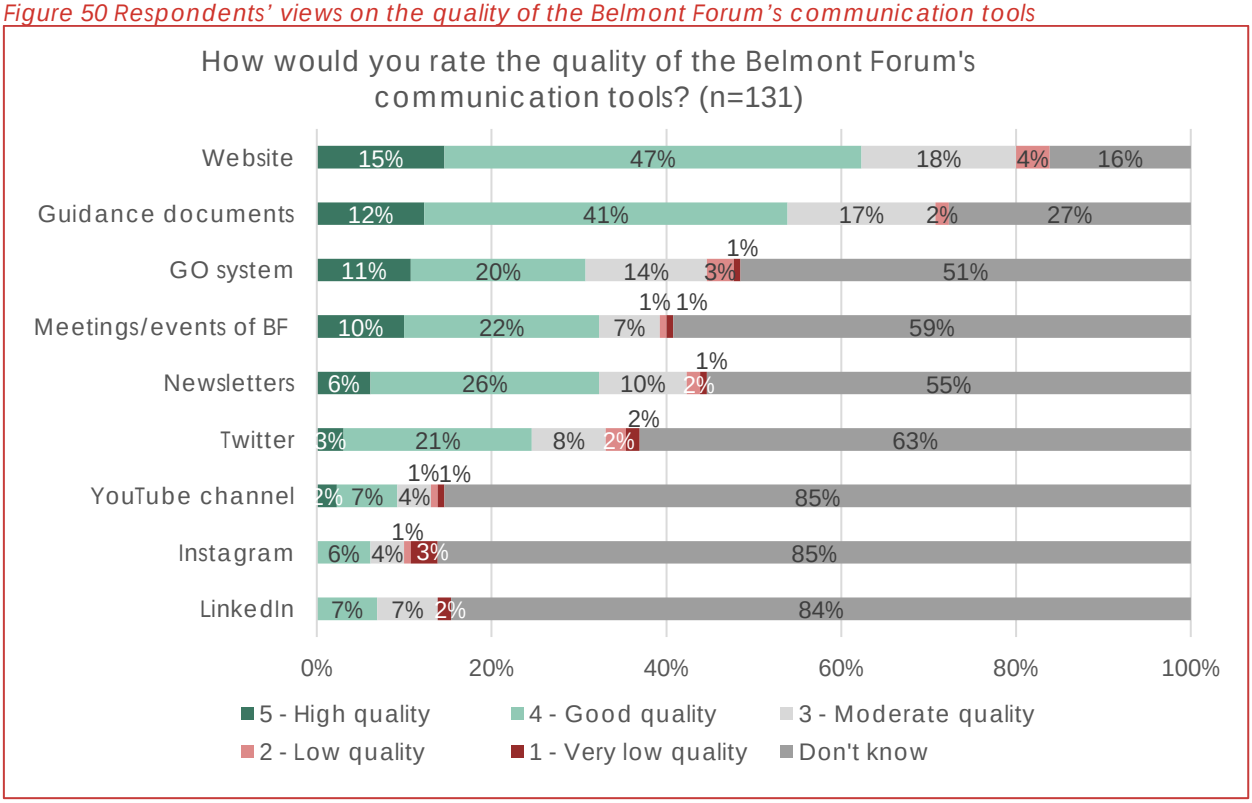
Figure 49 Respondents' views on the scientific impact of their projects



A.5.6 *Section F: Communication and dissemination*

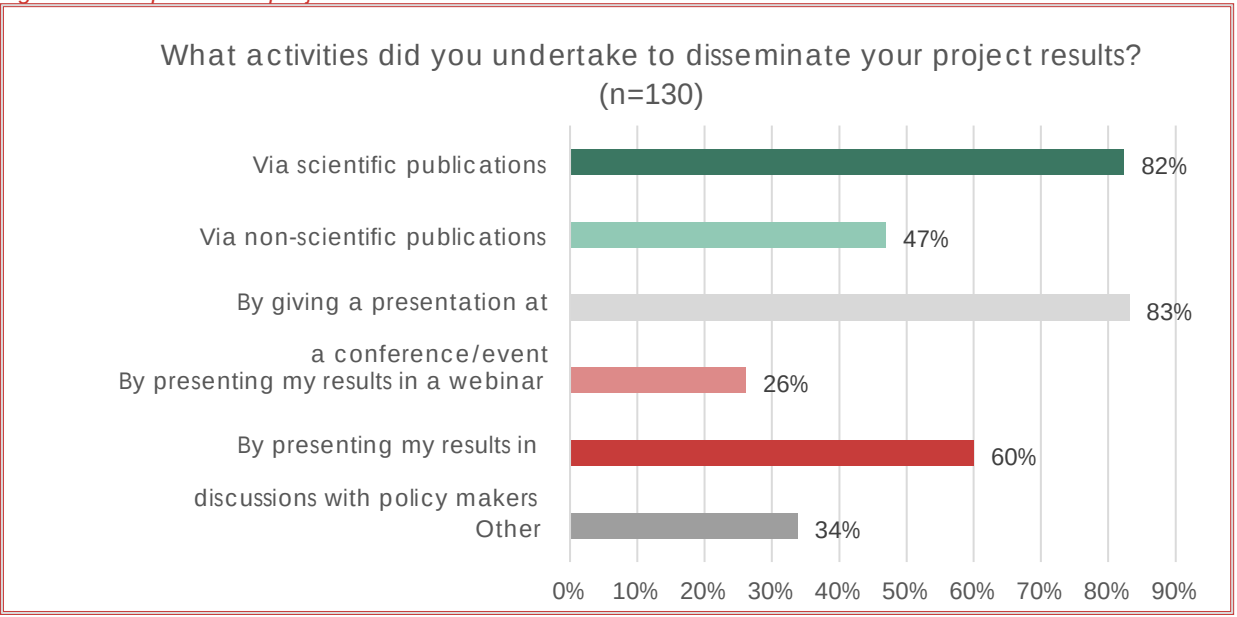
Overall, respondents rate the quality of the Forum's communication tools positively (in the majority good to high quality); however, the awareness level for most of them is low. The Forum's website and its guidance documents are the most familiar tools to the respondents, the social media channels are however least known. The quality of the Forum's LinkedIn communication tool is assessed least on quality. The low awareness levels may be associated with the fact that many of the respondents have just recently entered in cooperation with the Form and have thus not yet had a chance to assess or utilise the communication tools.

Figure 50 Respondents' views on the quality of the Belmont Forum's communication tools



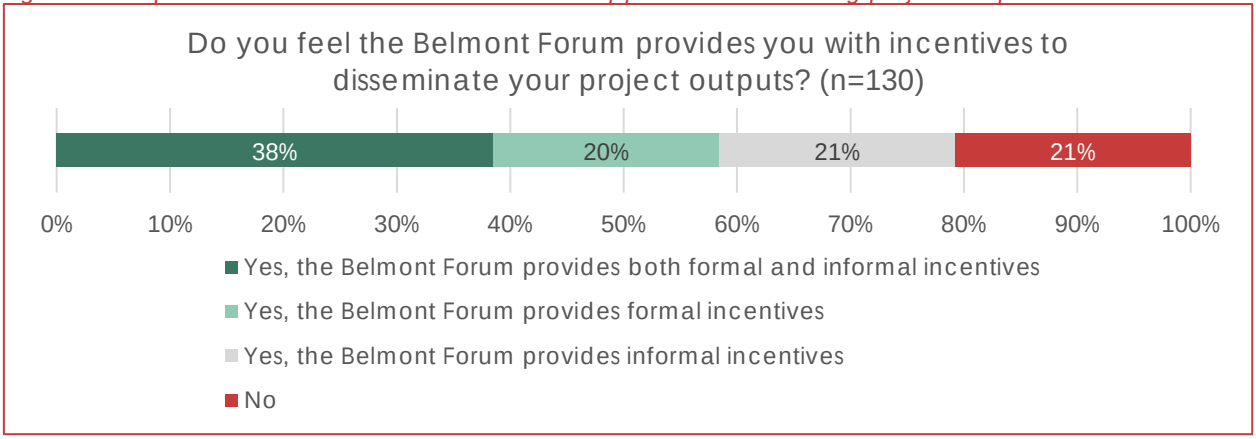
Various project dissemination activities have been used in projects. Results have most often been presented at conferences or events, followed by scientific publications and by presenting results in discussion with policymakers. Webinars have been undertaken the least.

Figure 51 Respondents project dissemination activities



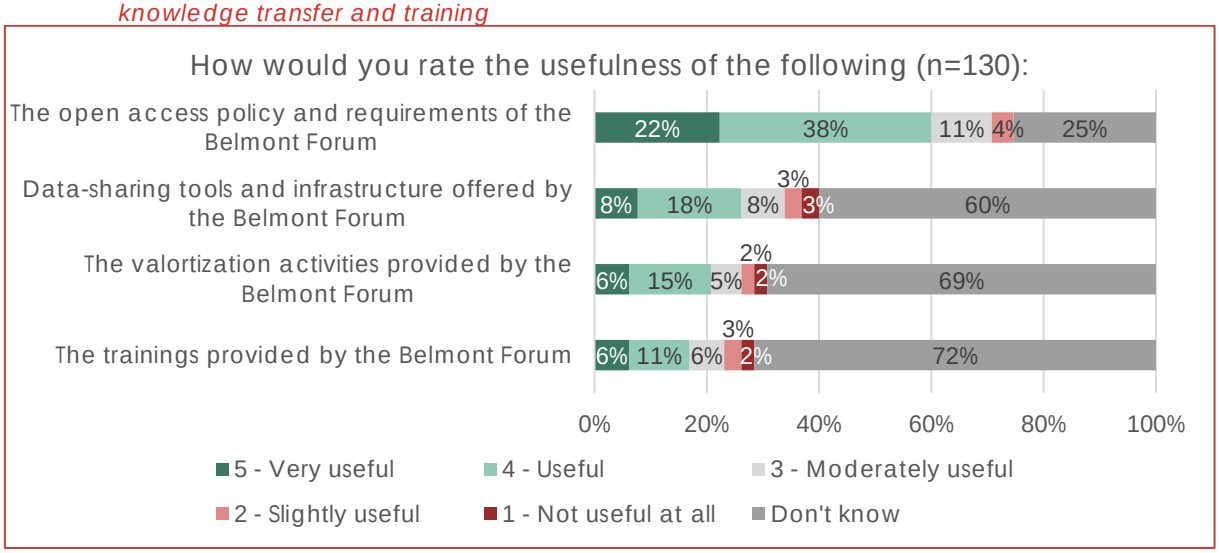
The majority of respondents believe that the Forum provides formal and informal incentives for the dissemination of project results; however, a small proportion (21%) disagrees.

Figure 52 Respondents' views on Belmont Forum's support in disseminating project outputs



In terms of Forum's tools for knowledge transfer and training, the majority of respondents are not sufficiently aware of those to make an informed selection. However, those who are informed rate the tools and trainings in the majority as useful or very useful.

Figure 53 Respondents' views on the usefulness of Belmont Forum's tools and support mechanisms for knowledge transfer and training



Appendix C Survey of Members and Partners of the Belmont Forum

The members and Partners of the Forum were surveyed in March-May 2020. The online survey was distributed to the members and Partners by the Forum Secretariat via a closed link (direct invitation) and was open for responses for 11 weeks. Survey invitations and reminders were sent via the Secretariat of the Forum. It was specified to the recipients to only provide one answer per organisation.

In total, we have received 35 responses to the survey. After removing (in consultation with the Forum) insufficiently filled partial responses and duplicate responses from identical organisations, 18 remaining responses were analysed using Microsoft Excel. In total, approximately half of the Members and Partners of the Forum have participated in the survey.

In the remainder of this appendix, we first provide an overview of the survey questions followed by an overview of the results of closed questions with annotated graphs.

A.6 Survey questions

Topic	#	Survey Question	Answer options
Respondent characteristics	A1	Are you a member or partner of the Belmont Forum?	Select from: Member of the Belmont Forum; Partner of the Belmont Forum; Partner in specific CRA's; Boundary organisation
	A2	Are you a member of the Steering Committee?	Select from: Yes, I am currently a member of the Steering Committee; No, but I was a member of the Steering Committee in the past; No, I have never been a member of the Steering Committee
	A3	What country are you based in?	Select from the list of world countries
	A4	What type of organisation do you work for?	Select from: Funding Agency; Ministry for science and research; Academy of Science; Scientific Institution; Other
	A5	Since when have you been a member of the Belmont Forum?	Select from: 2009-2020
	A6	Since when have you been a partner of the Belmont Forum?	Select from: 2009-2020
Mission and objectives	B1	In your opinion, which of the following descriptions represents the main objective of the Belmont Forum best? Please rank them in order of importance. • To fund multinational interdisciplinary projects that generate scientific results to better understand, mitigate, and adapt to global environmental change • To fund multinational transdisciplinary scientific projects that generate scientific research as well as interact with stakeholders	Rank the options in order of importance.

		in order to increase knowledge on mitigating and adapting to global environmental change • To fund multinational transdisciplinary scientific projects that are mainly aimed at achieving impact with stakeholders related to mitigating and adapting to global environmental change	
	B2	What are your organisation's main reasons to participate in the Belmont Forum? Please rank the importance of various potential reasons below • To fund scientific projects in the area of global environmental change that otherwise could not be realised • To realise international cooperation that otherwise could not be realised • To increase attention for and funding of transdisciplinary projects in the field of mitigating and adapting to global environmental change • To get or stay informed on international discussions around topics related to global environmental change • To contribute to and help steer international discussions around topics related to global environmental change • To meet, discuss, learn from, and exchange best practices with: funding agencies and/or organisations with similar objectives from other countries	Select from: 1 - not Important at all; 2 - slightly important; 3 - moderately important; 4 - important; 5 - very important; Not applicable
	B3	Other, namely:	Open answer box
	B4	On a scale from 1-10, how satisfied are you with: • The Belmont Forum in general • Your influence on the strategic direction taken by the Belmont Forum	Select from: 1- Highly dissatisfied – 10 - Highly satisfied
Design and implementation of CRAs	C1	How satisfied are you with the scoping process for CRAs? • Discussions in the plenary meetings on proposed topics • The use of scoping workshops as a method to develop CRA's • Number of scoping workshops • Depth of scoping workshops • Level of participation and inclusiveness in scoping workshops	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	C2	Please elaborate how could the scoping process be improved in the future:	Open answer box
	C3	How satisfied are you with the chosen CRAs?	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	C4	To what extent have CRAs overlapped with other initiatives or funding mechanisms?	Select from: 1 - not at all; 2 - to a small extent; 3 - to some extent; 4 –

			to a large extent; 5 - to a very large extent; Not applicable; Don't know
Day-to-day operations of the Belmont Forum	D1	To what extent are you satisfied with the internal coordination and communication between members? • Quality of interactions between the members • Frequency of interactions between the members (annual plenary meeting and possibly online contacts)	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D2	What could be improved in the future regarding the communication between members?	Open answer box
	D3	To what extent are you satisfied with the internal coordination and communication between members and the secretariat? • Quality of interactions with the secretariat • Frequency of interactions with the secretariat • Level of responsiveness of the secretariat	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D4	What could be improved in the future regarding communication between members and the secretariat?	Open answer box
	D5	How satisfied are you with the Belmont Forum's plenary meetings? • Frequency of plenary meetings • Agenda-setting process before the meetings • Provision of relevant information before the meetings • Quality of discussions during the meetings • Reporting after the meeting • Follow-up of action points after the meeting	Select from:1 – very dissatisfied; 2 – dissatisfied; 3 – somewhat satisfied; 4 – satisfied; 5 – very satisfied; Don't know
	D6	Please elaborate on how could the plenary meetings be improved in the future:	Open answer box
	D7	Do you feel that the governance, management, and organisation of the Forum continues to be adequate in the future (in view of the growth of the Forum)?	Open answer box
	D8	How satisfied are you with the role of the Steering Committee? Please indicate to what extent you agree with the following statements. • The tasks and responsibilities of the Steering Committee are well-defined and clear • The Steering Committee has sufficient mandate to play a meaningful role in the governance of the Belmont Forum • The Steering Committee is informed appropriately by the secretariat on crucial information regarding the (financial and other) performance of the Belmont Forum	Select from: 1 - strongly disagree; 2 - disagree; 3 - somewhat agree; 4 – agree; 5 - strongly agree; Don't know

		• The Steering Committee feeds back information adequately to the other members of the Belmont Forum The Steering Committee takes feedback from the other members sufficiently into account in its decision-making processes	
Assessment of the Forum function	E1	How satisfied are you with the number of members that are currently involved in the Belmont Forum? The current number of members (31 in total) is…	Select from: Too high; Appropriate; Too low
	E2	How satisfied are you with the variety of members currently involved in the Belmont Forum? • The geographical spread is • The diversity in types of members is	Select from: Too large; Appropriate; Too limited; Don't know
Resources of the Belmont Forum	F1	How satisfied are you with the contributions of members to the Belmont Forum? • The direct financial contributions are • The in-kind contributions are	Select from: Too large; Appropriate; Too limited; Don't know
	F2	To what extent do you feel that the current funding modalities are sufficiently flexible to attract new members? The current range of funding modalities are…	Select from: Too extensive (there is a need for more consistency); Sufficient (they are pragmatic for the BF going forward); Too limited (there is a need to offer more flexibility); Don't know
The Belmont Forum going forward	G1	To what extent do you agree with the following statement? "The Belmont Forum has been able to build up a strong and sustainable network of global change research funders and other relevant partners."	Select from: 1 - strongly disagree; 2 - disagree; 3 - somewhat agree; 4 – agree; 5 - strongly agree; Don't know
	G2	To what extent have the following added value(s) been achieved by the Belmont Forum? "The Belmont Forum has led to…" • Leverage of (national) investments • More interaction between global change research • funders/network building • Enhanced interdisciplinary knowledge • Enhanced opportunities for participation by researchers in Global South	Select from: 1 - strongly disagree; 2 - disagree; 3 - somewhat agree; 4 – agree; 5 - strongly agree; Don't know
	G3	Please specify any other added-value(s) below:	Open answer box
The Belmont Forum going forward	H1	What would be your advice for the Belmont Forum going forward?	Open answer box

A.7 Survey results

This section presents a basic analysis of the closed (multiple-choice) questions in the survey of Members and Partners of the Belmont Forum. Responses to open questions are not presented. Questions are presented in the order as outlined in the table above. Brief descriptions are provided.

A.7.1 *Section A: Respondent characteristics*

Of 18 received responses received to the Member/Partners survey, only one indicated being a partner of the Forum. The majority of the respondents are or have formerly been a member of the Steering Committee (Figure 25).

Figure 54 Responses per type of involvement with the Belmont Forum

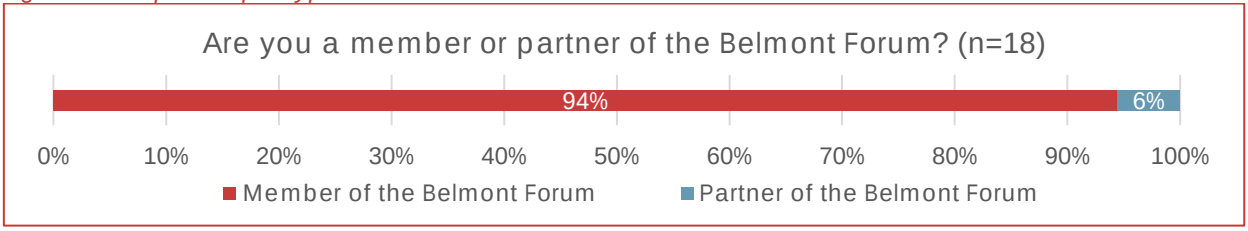
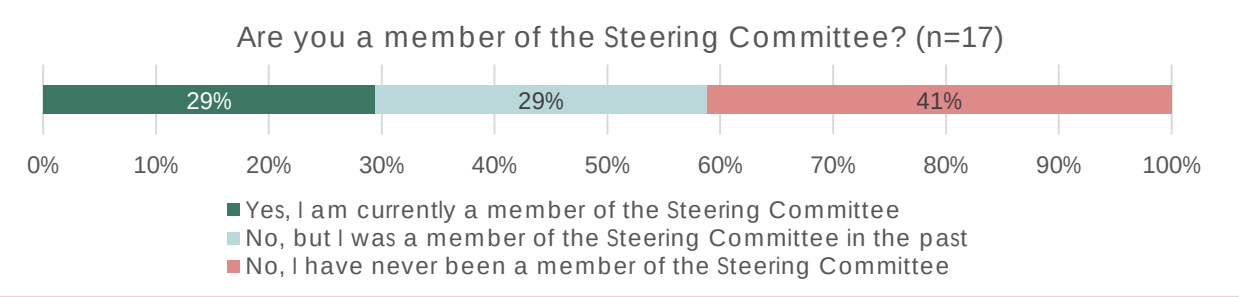
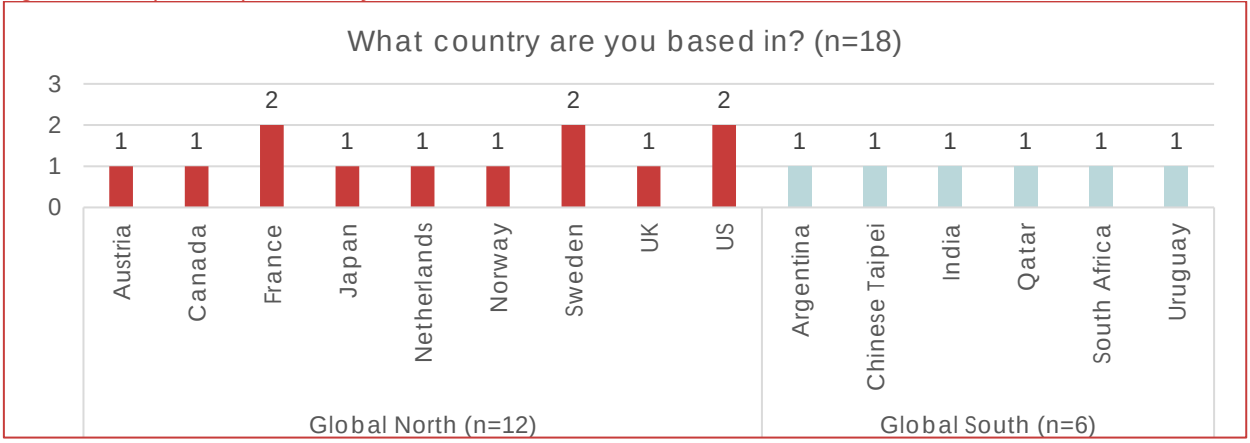


Figure 55 Responses per Steering Committee membership



In terms of country representation, the majority of respondents are from the Global North. The United States are represented by two organisations, one of which is a partner of the Forum.

Figure 56 Responses per country



The majority of respondents are from the funding agencies and the national ministries. There is a balanced distribution of responses in terms of length of involvement with the Forum (Figure 28), ranging from less than a year to over 10 years.

Figure 57 Responses per type of organisation

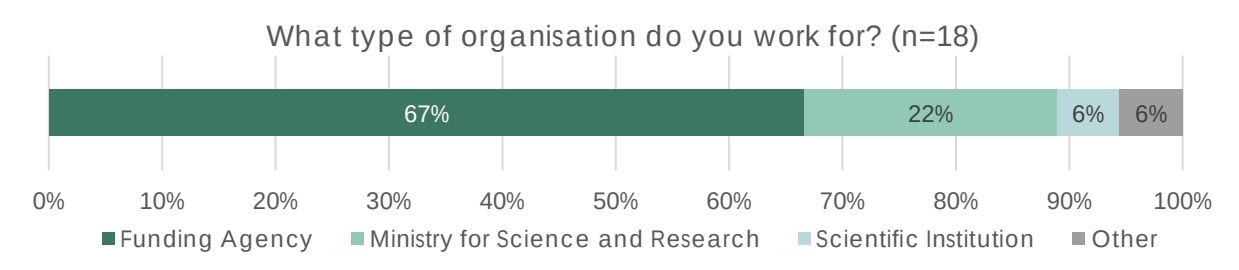
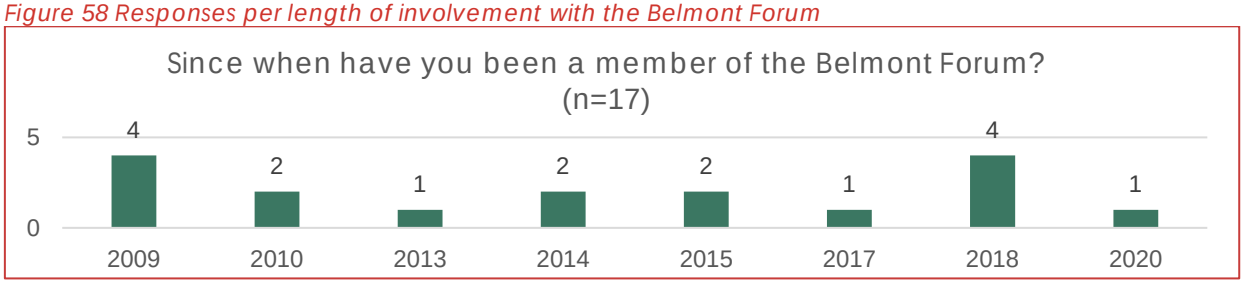


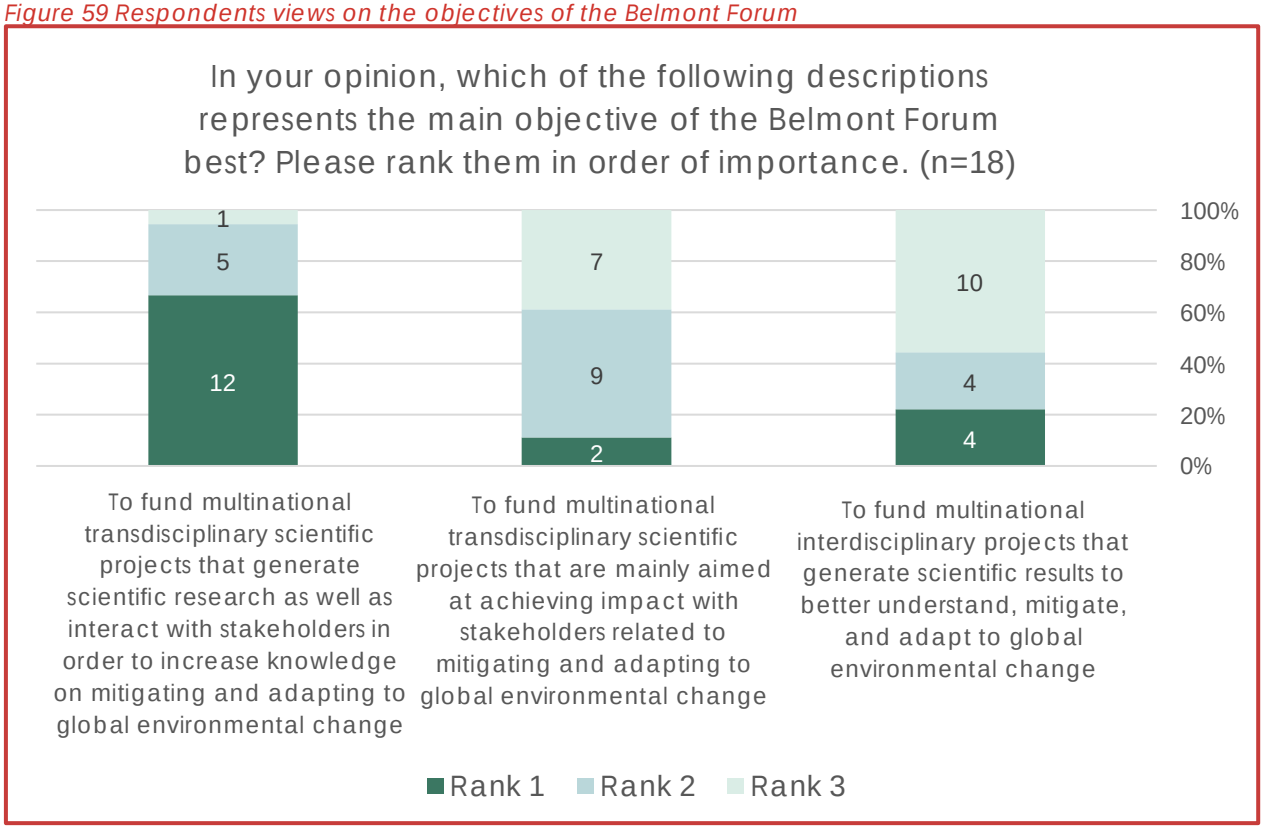
Figure 58 Responses per length of involvement with the Belmont Forum



A.7.2 *Section B: Mission and objectives*

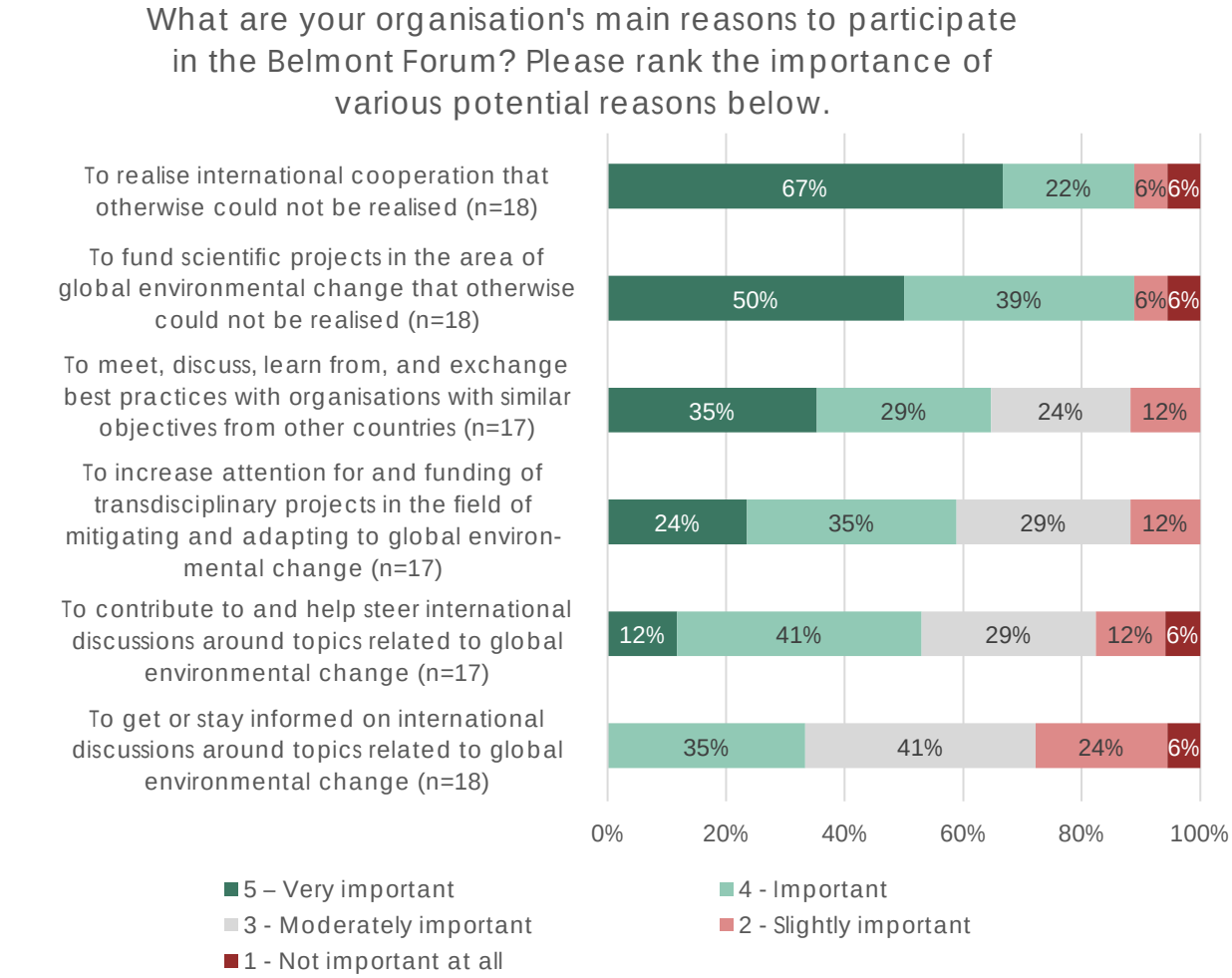
Overall, when ranking the objectives of the Forum, respondents value the stakeholder interaction aspect highly. The top-ranked option for the objective of the Forum is: "To fund multinational transdisciplinary scientific projects that generate scientific research as well as interact with stakeholders to increase knowledge on mitigating and adapting to global environmental change".

Figure 59 Respondents views on the objectives of the Belmont Forum



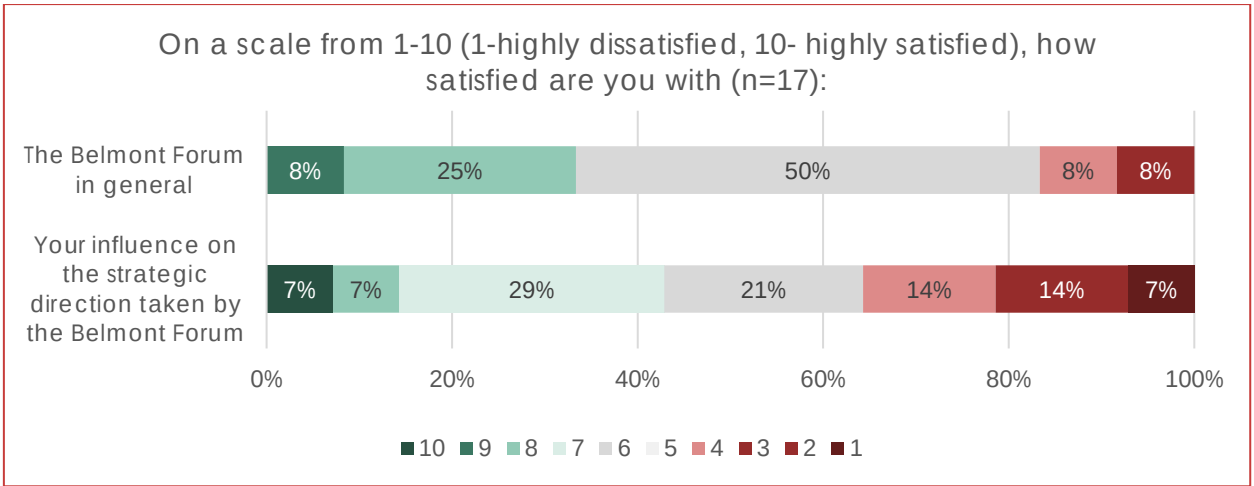
Of the various reasons to participate in the Forum, international cooperation, and opportunity to realise unique scientific research are top reasons considered important by the respondents.

Figure 60 Respondents reasons to participate in the Belmont Forum



The overall satisfaction with the Forum is somewhat low among the respondents: about 50% express only moderate levels of satisfaction. However, over 40% of the respondents are generally satisfied with their influence on the strategic direction taken by the Forum.

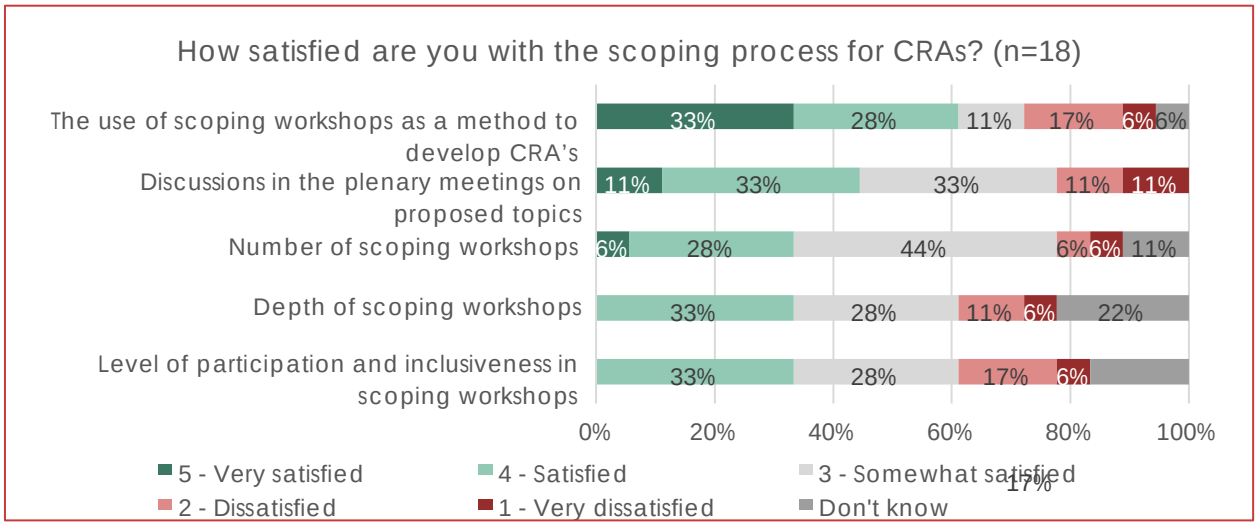
Figure 61 Respondents overall satisfaction with the Belmont forum and their participation in its strategic direction



A.7.3 *Section C: Design and implementation of CRAs*

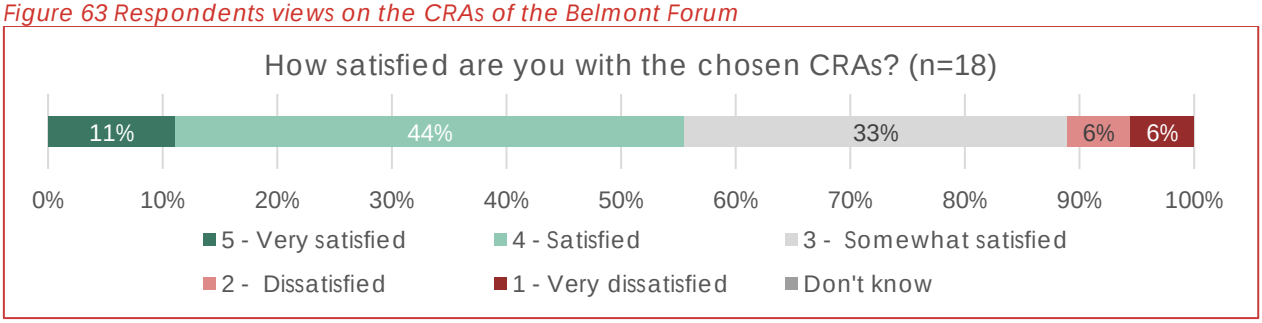
In general, respondents are satisfied with the use of scoping workshops for CRA development. However, less than half of the respondents feel highly satisfied with the implementation of the CRA scoping process (number, depth, level of participation and inclusiveness). It should be however noted that as there have been multiple scoping events, satisfaction rates may vary per event, which has not been asked in this survey.

Figure 62 Respondents views on the CRA scoping process



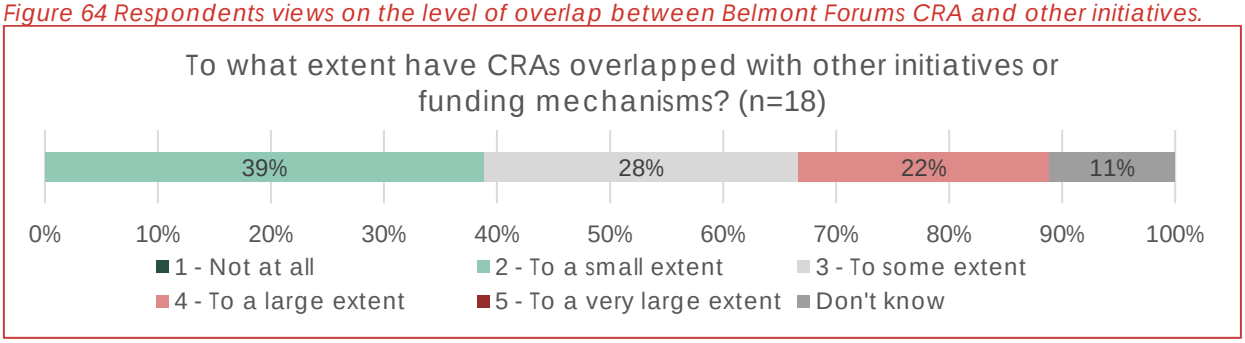
The level of satisfaction with the chosen CRAs varies among the respondents: over half are satisfied, while the rest indicate various levels of (dis)satisfaction.

Figure 63 Respondents views on the CRAs of the Belmont Forum



All participants indicated that there is a certain level of overlap between the CRA's of the Forum and other initiatives and funding mechanisms for research.

Figure 64 Respondents views on the level of overlap between Belmont Forums CRA and other initiatives.



A.7.4 *Section D: Day-to-day operations of the Belmont Forum*

Respondents expressed generally little satisfaction with internal coordination and communication between the members of the Forum. Higher satisfaction rates are reported for the internal coordination and communication between the members and the Secretariat, although the overall satisfaction with the quality of interaction between them is low (Figure 36).

Figure 65 Respondents' satisfaction with internal coordination between the members of the Belmont Forum

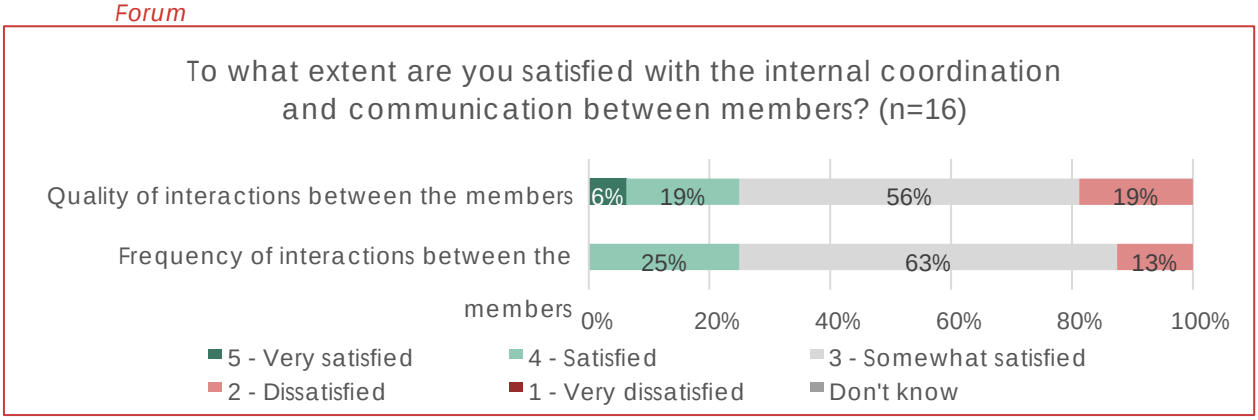
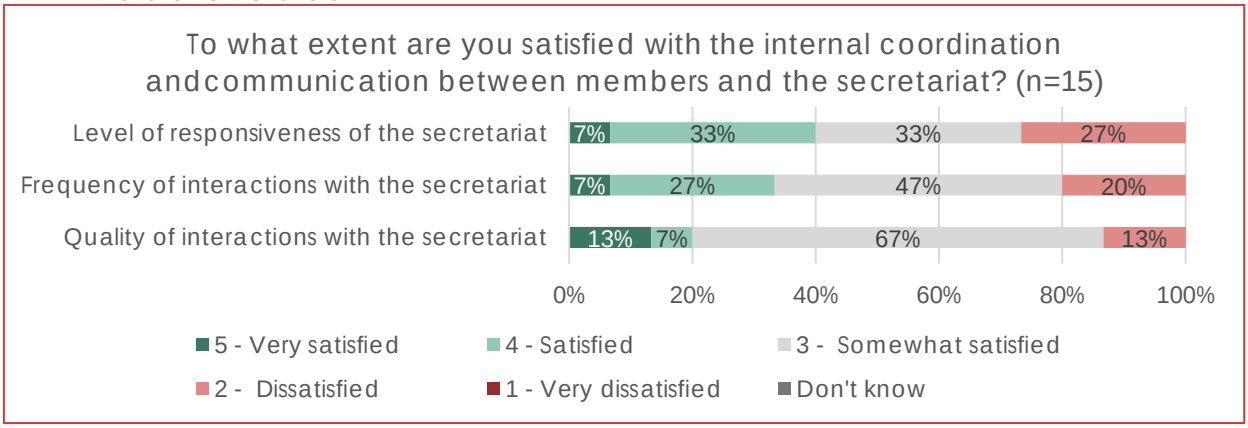
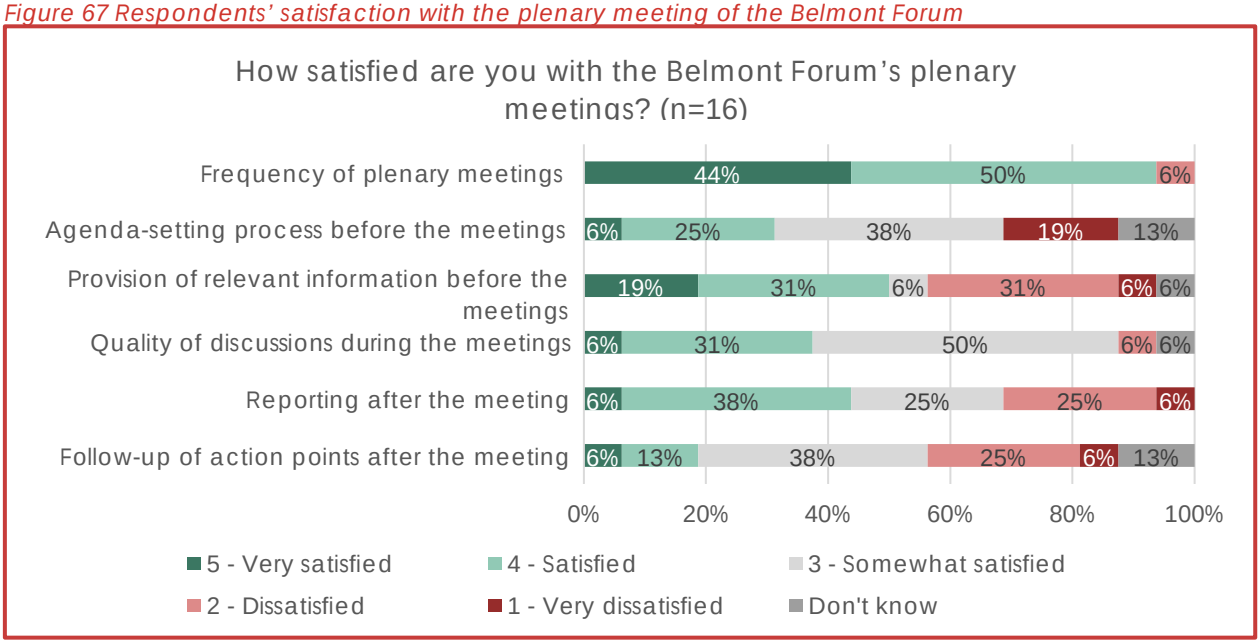


Figure 66 Respondents' satisfaction with internal coordination between the members and the Secretariat of the Belmont Forum



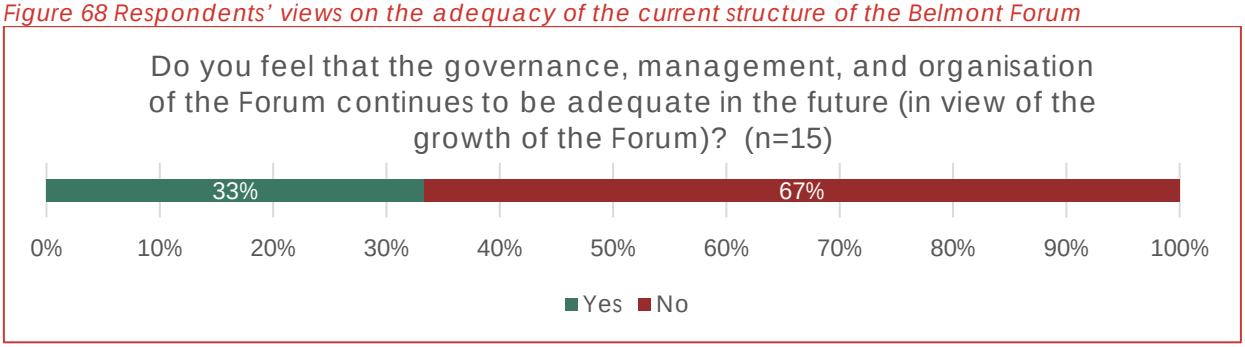
Respondents are generally satisfied with the Forum's plenary meetings, however, satisfaction with logistical aspects, particularly with agenda-setting and follow up processes, is lower and clearly reveals opposing views.

Figure 67 Respondents' satisfaction with the plenary meeting of the Belmont Forum



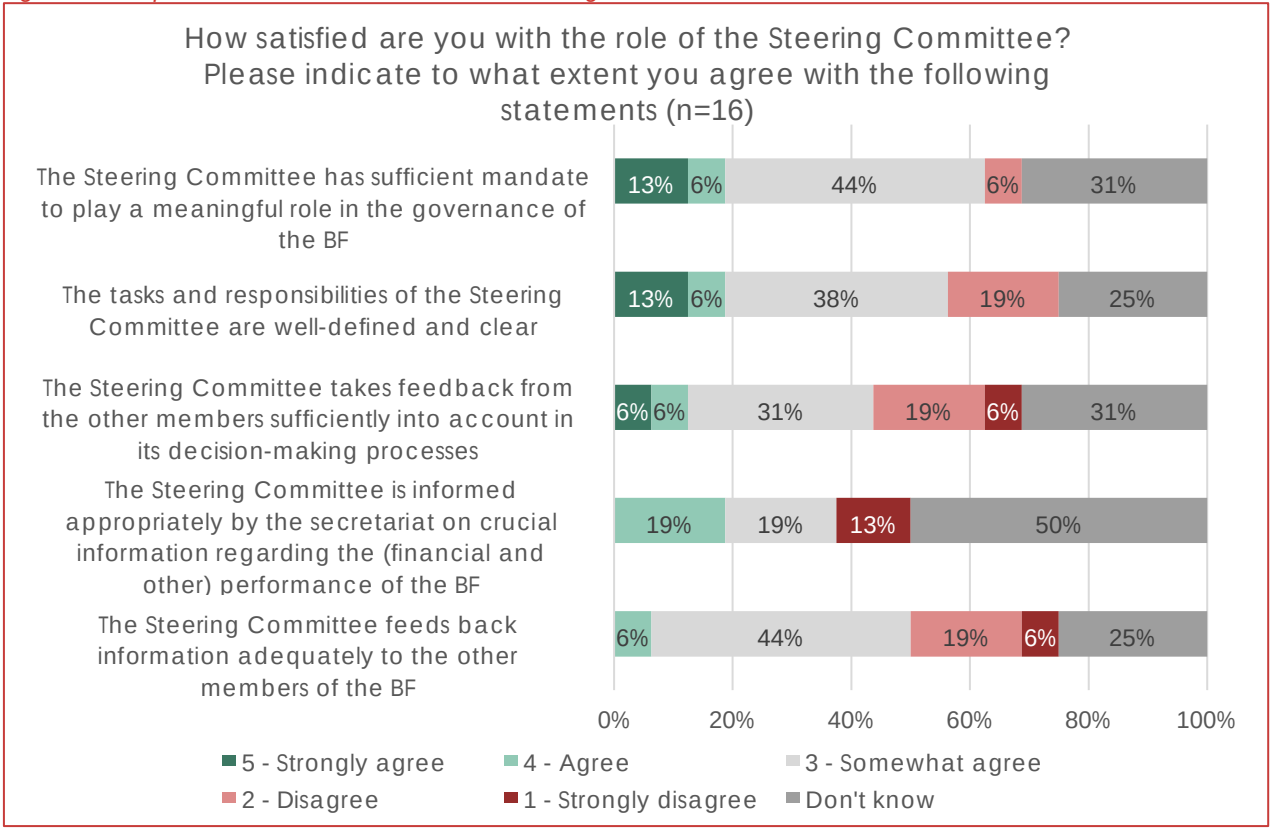
Most of the respondents believe that the current governance, management, and organisation of the Forum is inadequate given the future growth of the organisation.

Figure 68 Respondents' views on the adequacy of the current structure of the Belmont Forum



The level of satisfaction with the role of the Steering Committee is neutral/low among the respondents. Many respondents did not answer this question, which raises concern over the extent of their familiarity with the role of the Steering Committee or their willingness to provide this type of feedback. It is also possible that those who did not provide answers are, in fact, members of the SC, thus they felt uncomfortable rating their own performance.

Figure 69 Respondents satisfaction with the Steering Committee's role



A.7.5 *Section E: Assessment of the Forum function*

The views on the current number of members of the Forum are split: over half of the respondents are satisfied with the current number of members, while one third find it too low. There is more consensuses in views on the variety of members (Figure 41): the graphical spread of members is considered too limited by most respondents, while the respondents are mostly content with the diversity in types (of organisations) of members.

Figure 70 Respondents satisfaction with the size of membership of the Belmont Forum

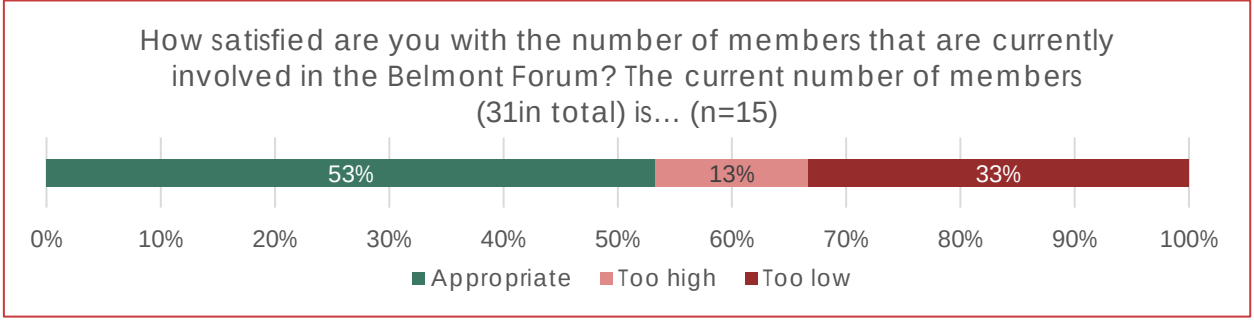
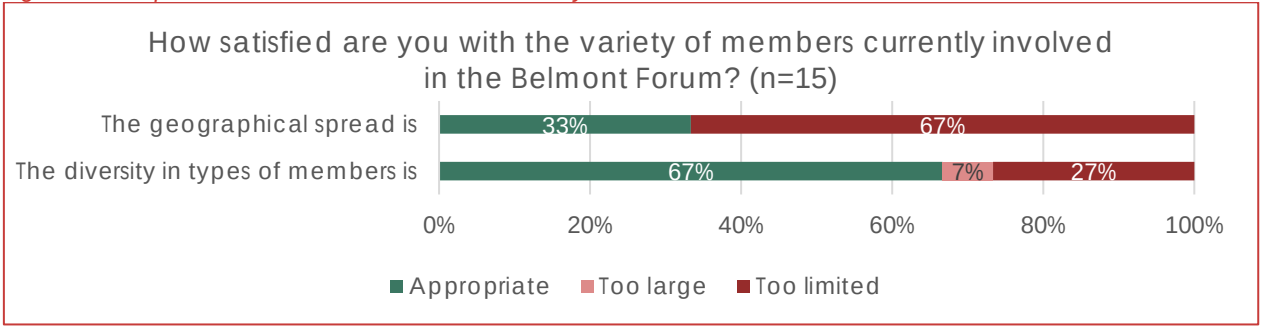


Figure 71 Respondents satisfaction with the variety of members involved in the Belmont Forum



A.7.6 *Section F: Resources of the Belmont Forum*

Respondents are generally satisfied with the size and flexibility (Figure 43) of member contributions, however about one third find both direct and in-kind contributions too low.

Figure 72 Respondents satisfaction with the size of member contributions

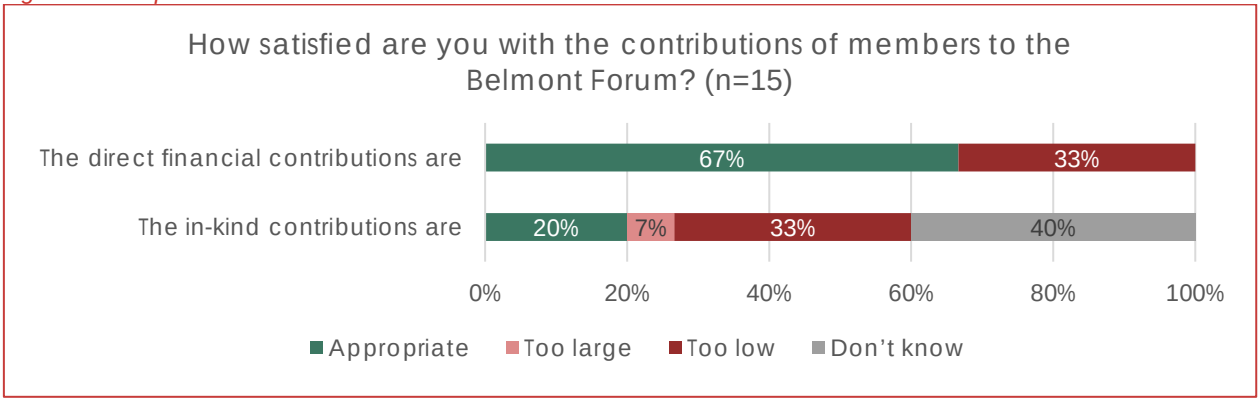
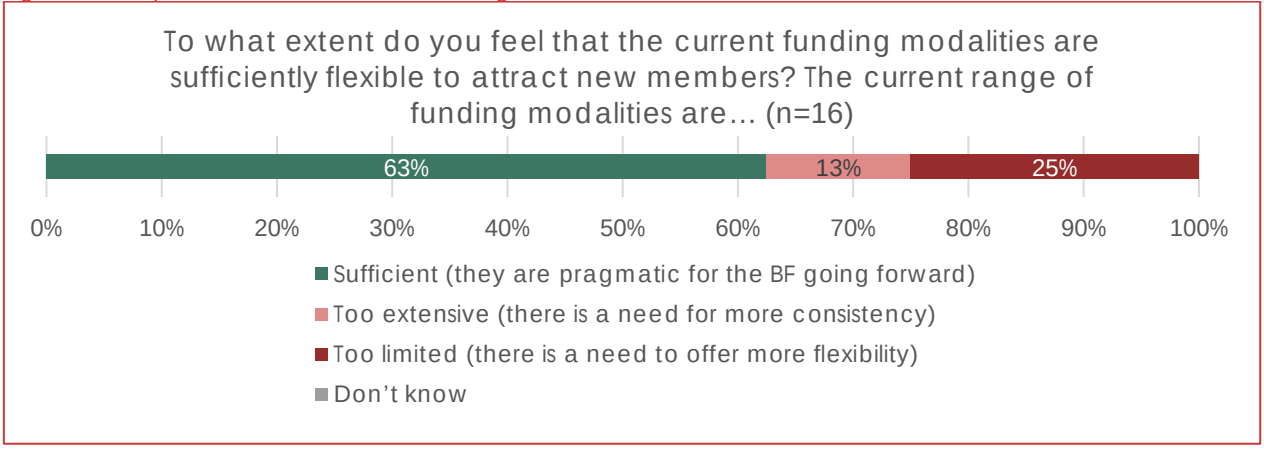


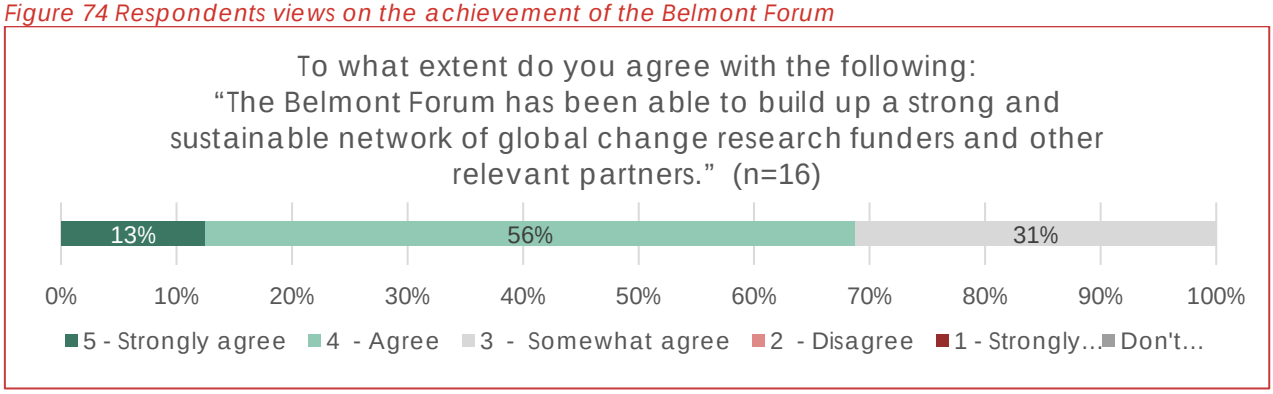
Figure 73 Respondents views on the funding modalities



A.7.7 *Section G: The Belmont Forum going forward*

Respondents generally agree that the Forum has been able to build a network of global change research funders and other relevant partners.

Figure 74 Respondents views on the achievement of the Belmont Forum



Appendix D Interview guides

To facilitate the involvement of all stakeholders in the evaluation process, face-to-face and phone in-depth interviews were conducted with representatives of the Secretariat, the members and Partners, and the beneficiaries of the Forum as well as its external stakeholders. Table 1 presents an overview of all conducted interviews. In total, 40 interviews were conducted including:

- 3 interviews with the members of the Secretariat of the Forum
- 20 interviews with the Members of the Forum
- 6 interviews with the Partners of the Forum
- 5 interviews with external stakeholders
- 6 interviews with the beneficiaries of the Forum

The majority of the interviews were conducted in a virtual setting; however, a number were held as physical meetings during the 10th Anniversary Meeting and Plenary in Taipei in October 2019.

The interviews were transcribed but not recorded. The analysis was performed using the qualitative analysis software ATLAS.ti, where gathered data was coded and grouped, and further analysed thematically according to the evaluation domains.

Further in this Annex we present interview guides with interview questions used for various stakeholder groups, namely:

- Interview guide for the Members and Partners of the Forum
- Interview guide for the staff and volunteers of the Forum
- Interview guide for external stakeholder of the Forum
- Interview guide for the beneficiaries of the Forum

Table 13 Overview of conducted interviews

Name	Organisation	Country	Stakeholder type
Erica Key	Belmont Forum	Global	Secretariat
Roel Marsman	European Commission	EU	Secretariat
Judit Ungvari	AAAS	USA	Secretariat
Marcella Ohira	IAI	Americas	Member, Host of Secretariat
Marcos Regis da Silva	IAI	Americas	Member, Host of Secretariat
Maria Uhle	NSF	USA	Member
Kate Hamer	NERC	UK	Member
Marialuisa Tamborra	European Commission	EU	Member
Minn-Tsong Lin	MoST	Chinese Taipei	Member
Harald Leisch	DFG	Germany	Member
Reynaldo Victoria	FAPESP	Brazil	Member
Stephanie Thiebault	AllEnvi	France	Member

Gansen Pillay	NRF	South Africa	Member
Dick van der Kroef, Maya Rispens	NWO	Netherlands	Member
Elizabeth Boston	NSERC	Canada	Member
Andrew Kaniki	NRF	South Africa	Member
Parvinder Maini	MOES	India	Member
Ipek Erzi	TUBITAK	Turkey	Member
Magnus Tannerfeldt	FORMAS	Sweden	Member
Yoshiko Shirokizawa	JST	Japan	Member
Omar Boukharis	QNRF	Qatar	Member
Yao Yupeng	NSFC	China	Member
Antonello Provenzale	CNR	Italy	Member
Michael Kuperberg	USGCRP	USA	Partner
Yaroslav Sorokotyaga	RFBR	Russia	Partner
Gilberto Câmara	GEO	n/a, international	Partner
Heide Hackmann	ISC	n/a, international	Partner
Albert van Jaarsveld	IIASA, NRF	n/a, international	Partner
Jon Padgham	START	USA	Partner
Patrick Monfray	Ministry of Research	France	External Stakeholder
Josh Tewksbury	Future Earth	USA	External Stakeholder
Thorsten Kiefer	JPI Oceans	Europe	External Stakeholder
Eleanor Robson	Future Earth Australia	Australia	External Stakeholder
Gen Tsukada	Asia Pacific Network	n/a, international	External Stakeholder
Beneficiary A	University	India	Beneficiary
Beneficiary B	University	Argentina	Beneficiary
Beneficiary C	University	United Kingdom	Beneficiary
Beneficiary D	University	Sweden	Beneficiary
Beneficiary E	Research institute	Australia	Beneficiary
Beneficiary F	University	USA	Beneficiary

Names of beneficiaries are not disclosed for confidentiality reasons

A.8 Interview guide for the members and partners of the Belmont Forum

Name interviewee	
Role	
Name interviewer	
Date of the interview	

Interviewer to introduce the interim review and the main purpose of this interview.

Introduction

1. Could you **briefly introduce yourself** and your organisation?
    - How long has your organisation been a member of the Belmont Forum?
    - What is your role within the organisation?

Relevance and mission

2. What were the main **reasons for your organisation to become a member/partner** of the Belmont Forum? Please elaborate.
3. What do you consider to be the **main problems and issues** in the area of global environmental change? Do you feel that the Belmont Forum's challenge and objectives address these problems and issues appropriately?
4. To what extent do you feel that the Forum should adjust its **scope, mission, and/or objectives in the future**? Can you explain?

Organisation

5. Do you feel that the Forum's **main processes and procedures are clear and appropriate** (scoping of CRAs, application, evaluation, selecting, and granting of projects)? In terms of:
    - The level of transparency and inclusiveness of procedures
    - The level of professionalism, responsiveness, and flexibility towards members, partners, applicants, and beneficiaries
6. Do you feel that the governance, management, and organisation of the Forum continues to be **adequate in the future** (in view of the growth of the Forum)? If not, what adjustments would be needed going forward?

Effectiveness and impact

7. To what extent are you satisfied with the way in which the **Belmont Forum engages with its members and partners**? How satisfied are you with the communications and interactions with the Belmont Forum? Please explain.
8. Do you feel that the Forum has been able to build a **strong and sustainable network** of global change research funders? Can you explain?

9. To what extent would your organisation (likely) **collaborate with the members and partners in other countries without the Forum**? In what way(s)?
10. To what extent do you feel that the Belmont Forum successfully stimulates researchers to truly ***co-design*** **projects with stakeholders in other countries, disciplines or sectors**?
11. What do you consider to be the **main results and impacts of the Belmont Forum** in the following areas? What are the main outputs? What is the actual impact? Please explain and use examples where possible?
    - Scientific impact
    - Wider dissemination of knowledge and information (esp. non-scientific outputs) to various audiences
    - Impact on policy making (esp. awareness and uptake among policymakers)
    - Other
12. To what extent are you **satisfied with these effects and impacts**?
    - How would you assess the usefulness of BF funded-projects and their outputs?
    - What could the Forum do to further enhance its impact?
13. How do you assess the **visibility and reputation** of the Belmont Forum? Please explain. What could be done to further raise awareness of the Forum around the globe?
14. Looking at the bigger picture, how would you assess the **Belmont Forum's overall impact**? What role does it play in addressing the issue of global environmental change? Especially:
    - Understanding
    - Mitigating
    - Adapting to global environmental change

    Note to interviewer: try to identify the main impact pathways

Efficiency

15. [*Note to interviewer: Important question]* How satisfied are you with the efficiency of the Belmont Forum's **internal processes and procedures**? Think of:
    - Scoping process CRAs
    - Launch of calls
    - Selection of panel experts
    - Application, evaluation, selection, awarding processes
    - Organisation/agenda-setting of plenary meetings
    - Any other topics related to efficiency
16. To what extent do you feel that the Belmont Forum is equipped with adequate **human resources (capacity and competences) and financial resources** to run the Forum smoothly and efficiently? Please explain.

Coherence

17. [*Note to interviewer: Important question]* To what extent do you feel that the selection of topics for the **CRAs as well as selected projects are coherent and complementar**y to each other (e.g. no contradictions, overlaps, etc.)? Please explain.

18. To what extent do you feel that the Belmont Forum as an initiative is **complementary to other initiatives** in the in the same field (i.e. global environmental change research)?

- What are the most relevant initiatives in the field?
- How do their mission and objectives relate to those of the Belmont Forum?
- How often (if at all) does the Forum interact and collaborate with these other initiatives?

Added value

19. What do you consider to be the main added value(s) of the Belmont Forum (i.e. what would not have been possible to achieve without the Forum)? Examples:

- Leverage of (national) investments
- More interaction between global change research funders / network building
- Enhancement of interdisciplinary knowledge
- Other

Sustainability

20. [*Note to interviewer: Important question]* How **sustainable are the results achieved** (by projects and the Forum) so far?

- Are these outputs accessible and useable for others?
- does the Forum do in order to encourage the continued use of results?

21. How sustainable is the Belmont Forum in your view **in the next ten years**?

- How much funding has the Forum already secured for the next years?
- What is done by the Forum to maintain and enhance the support and investments from (existing and new) members and partners?

A.9 Interview guide for the staff and volunteers of the Belmont Forum

Name interviewee	
Role	
Name interviewer	
Date of the interview	

Interviewer to introduce the interim review and the main purpose of this interview.

Introduction

1. Could you briefly **introduce yourself**? What is your role in the Belmont Forum? How long have you been in this role/position?

Rationale and mission of the Belmont Forum

2. What was the main **rationale** for establishing the Belmont Forum ten years ago?
    - What were the main issues and problems observed in understanding, mitigating, and adapting to global environmental change?
    - Can you point us to any documentation discussing these issues and problems?
3. Do you feel that the **needs and problems have changed over time**? Can you explain?
4. To what extent do the rationale and mission of the Belmont Forum **continue to be relevant in the future**?
    - Is there a need for any adjustments in the way the mission and Challenge are defined?
    - Do you foresee any new or developing issues or needs in the future (think of scientific and policy developments in particular)?

Explanation of the Belmont Forum

5. Could you explain how the Forum is set up? What does the **financing system** look like?
6. What does the **scoping process of CRAs** look like? How are members involved in this?

Governance, management and organisation

7. Do you feel that the **division of roles and responsibilities** between the following bodies of the Forum is sufficiently clear? Please explain.
    - Secretariat
    - Secretariat host (IAI)
    - Steering committee
    - Group of program coordinators
    - Thematic program offices
8. How satisfied are you with the **interaction between your own and other parts of the Belmont Forum**?
    - How often are you in contact with your colleagues in the various bodies?

- What do these formal and informal contacts look like?

9. Do you feel that the Forum's **main processes and procedures are clear and appropriate** (scoping of CRAs, application, evaluation, selecting, and granting of projects, internal administrative procedures)? In terms of:
   - The level of transparency and inclusiveness of procedures
   - The level of professionalism, responsiveness, and flexibility towards members, partners, applicants, and beneficiaries
10. Do you feel that the governance, management, and organisation of the Forum continues to be **adequate in the future** (in view of the growth of the Forum)? If not, what adjustments would be needed going forward?
11. What (if any) mechanisms are currently in place to **monitor the progress and results** of the Forum? How effective are these mechanisms in your view?

Efficiency *[Note to interviewer: Important questions]*

12. Do you feel that you have **sufficient time, resources, and competences** to carry out your personal tasks and activities? Please elaborate.
13. To what extent are you satisfied with the efficiency of **internal coordination and communication**? What (if anything) would be needed to solve any issues in the future?
14. To what extent are you satisfied with the efficiency of the following **internal processes**?
    - Launch of calls
    - Selection of panel experts/evaluators
    - Organisation plenary Steering Committee meetings (preparation, agenda-setting, etc.)
    - Scoping of CRAs
    - Application, evaluation, selecting, and granting of projects

Communications and interactions with target audiences

15. How satisfied are you with the Forum's **interactions and engagement with its key target audiences** (members, partners, applicants, and beneficiaries)?
    - To what extent do you feel that the Forum has successfully built up strong and sustainable partnerships with its stakeholders?
    - Do you feel that the Forum is aware of the key issues that its stakeholders face, any struggles or questions that may have, etc.?

Progress and results

16. How satisfied are you with the progress and results achieved by the Forum so far? What do you consider the **main successes and achievements** in the last ten years? Please elaborate:
    - Network building
    - Scientific achievements
    - Dissemination of knowledge and information

- Policy achievements
- Other achievements

17. What were the **main challenges, obstacles or hindering factors**? Are there any areas where you would still like to improve things?
18. Does the Forum get **approached by policymakers** for advice or dialogue? If so, can you explain on what occasions, how often, by whom, for what purposes, etc.?

Visibility and dissemination

19. What (formal) requirements and (informal) incentives are used by the Belmont Forum to **encourage beneficiaries to disseminate their project-outputs**?
20. What does the **Belmont Forum do to disseminate** (promising) project-outputs itself?
21. How satisfied are you with the **visibility and reputation of the Forum** around the globe?
    - What is currently done to attract new and more resource funders?
    - How satisfied are you with the number and quality of proposals received by the Forum?
    - Do you observe any differences in visibility or reputation between countries or types of stakeholders?
    - Are there any specific outreach activities to (better) reach and include the Global South in the Belmont Forum?

Coherence

22. How satisfied are you with the **level of complementarity** between the Belmont Forum and other initiatives in the same field (i.e. global environmental change research)?
    - What are the most relevant initiatives in the field?
    - How do their mission and objectives relate to those of the Belmont Forum?
    - How often (if at all) does the Forum interact and collaborate with these other initiatives?

Sustainability

23. [*Note to interviewer: Important question]* How **sustainable are the results** achieved (by projects and the Forum) so far?
    - Are these outputs accessible and useable for others?
    - What does the Forum do in order to encourage the continued use of results?
24. How **sustainable is the Belmont Forum** in your view in the next ten years?
    - How much funding has the Forum already secured for the next years?
    - What is done by the Forum to maintain and enhance the support and investments from (existing and new) members and partners?

A.10 Interview guide for the external stakeholders of the Belmont Forum

Name interviewee	
Role	
Name interviewer	
Date of the interview	

Interviewer to introduce the interim review and the main purpose of this interview.

Introduction

1. Could you **briefly introduce yourself** and your organisation?
    - What is your role within the organisation?
    - How familiar are you with the Belmont Forum and its activities?

Rationale and mission

2. What do you consider to be the **main problems and issues** in the area of global environmental change?
    - Do you feel that the Belmont Forum's challenge and objectives address these problems and issues appropriately?
3. To what extent do you feel that the Forum should adjust its **scope, mission, and/or objectives in the future**? Can you explain?

Effectiveness and impact

4. Do you feel that the Forum has been able to build a **strong and sustainable network** of global change research funders? Can you explain?
5. How do you assess the **visibility and reputation** of the Belmont Forum? Please explain. What could be done to further raise awareness of the Forum around the globe?
6. What do you consider to be the **main results and impacts of the Belmont Forum** in the following areas? Please explain and use examples where possible?
    - Scientific impact
    - Wider dissemination of knowledge and information (esp. non-scientific outputs) to various audiences
    - Impact on policy making (esp. awareness and uptake among policymakers)
    - Other
7. How would you assess the **Belmont Forum's overall impact**? What role does it play in addressing the issue of global environmental change? Especially:
    - Understanding
    - Mitigating
    - Adapting to global environmental change
8. What could the Forum do to further enhance its impact?

Coherence

9. To what extent do you feel that the selection of topics for the **CRAs as well as selected projects are coherent and complementary** to each other (e.g. no contradictions, overlaps, etc.)? Please explain.

10. To what extent do you feel that the Belmont Forum as an initiative is **complementary to other initiatives** in the in the same field (i.e. global environmental change research)?

    - What are the most relevant initiatives in the field?
    - How do their mission and objectives relate to those of the Belmont Forum?
    - How often (if at all) does the Forum interact and collaborate with these other initiatives?

Added value

11. What do you consider to be the main added value(s) of the Belmont Forum (i.e. what would not have been possible to achieve without the Forum)? Examples:

    - Leverage of (national) investments
    - More interaction between global change research funders / network building
    - Enhancement of interdisciplinary knowledge
    - Other

Sustainability

12. How sustainable is the Belmont Forum in your view **in the next ten years**?

Conclusion

13. Is there anything else you would like to mention?

A.11 Interview guide for the beneficiaries of the Belmont Forum

Name interviewee	
Role	
Name interviewer	
Date of the interview	

Interviewer to introduce the interim review and the main purpose of this interview.

Interviewer to consult survey responses of interviewee before the start of the interview.

Introduction

1. Could you **briefly introduce yourself** and your organisation?
    - Is your organisation a member or a partner of the Belmont Forum?
    - What is your role within the organisation?
2. What **CRAs did you apply** for in the past? (*consult survey responses)*
    - Where these applications accepted/rejected?
3. [*If applicable*] Could you please elaborate on your **most recent project** funded by the Belmont Forum?
    - Are you project lead or participant in the project?
    - How did this project come about? (who came up with the idea, how did it evolve, why did you choose for certain partners, who designed the project, what was your role in the design, etc.)
4. [*Note to interviewer: important question]* The Belmont Forum aims for the organisation to support research projects that are "**co-developed and co-implemented** (...) with communities, policy makers, business and industry, ..."
    - To what extent do you feel that the Belmont Forum successfully stimulates researchers to truly ***co-design*** **projects with stakeholders in other countries, disciplines or sectors**?
    - To what extent do you feel that you have realised this ambition towards **stakeholder engagement** in your project? Can you provide an example?
    - To what extent are **North-South collaborations** realised in the core work of your project? Can you provide an example?

Relevance and mission

5. What do you consider to be the **main problems and issues** in the area of global environmental change?
    - Do you feel that the Belmont Forum's challenge and objectives address these problems and issues appropriately?

- According to the survey, you feel the activities of the Belmont Forum contribute to understanding, mitigating and adapting to global environmental change to a [*insert survey answer*] extent. **Why is this**?

6. To what extent do you feel that the Forum should adjust its **scope, mission, and/or objectives in the future**? Can you explain?

Organisational set-up & CRAs

7. [*check survey responses*] Do you feel that the Forum's **main processes and procedures are clear and appropriate** (scoping of CRAs, application, evaluation, selecting, granting of projects, mid-term meeting, final meeting)? In terms of:
   - The level of transparency and inclusiveness of procedures
   - The level of professionalism, responsiveness, and flexibility towards members, partners, applicants, and beneficiaries

Effectiveness and impact

8. [*Note to interviewer: important question]* What do you consider to be the **main results and impacts of the Belmont Forum** in the following areas? What are the main outputs? What is the actual impact? Please explain and use examples where possible?
   - Scientific impact
   - Wider dissemination of knowledge and information (esp. non-scientific outputs) to various audiences
   - Impact on policy making (esp. awareness and uptake among policymakers)
   - Other
9. To what extent are you **satisfied with these effects and impacts**?
   - How would you assess the usefulness of your project funded by the BF and its outputs?
   - What could the Forum do to further enhance its impact?
10. How do you assess the **visibility and reputation** of the Belmont Forum? Please explain. What could be done to further raise awareness of the Forum around the globe?

Coherence

11. To what extent do you feel that the Belmont Forum as an initiative is **complementary to other initiatives** in the in the same field (i.e. global environmental change research)?
    - What other relevant initiatives in the field you are familiar with?
    - At what other initiatives do you apply for project funding?
    - How do these initiatives relate to the Belmont Forum?

Added value

Note to interviewer: important question, ensure to elaborate

12. Please elaborate on the reason why you rate:

- The likeliness of your **current research collaboration** as [*very likely/not likely*] without funding from the Belmont Forum
- The likeliness of **continued existence of your project/ continued collaboration** with your consortium partners as [*very likely/not likely*] after funding from the Belmont Forum ends
    - If likely: In what way and form? Are there concrete ideas for continuance? With the same consortium?
    - If not likely: why not?

Communication and dissemination

13. [*Note to interviewer: Important question]* What were the **main types of non-journal outcomes from your project**?

- Did you document these outcomes in your annual reports and/or a project website?
- Is there any public source available that provides a description of these outcomes?

14. How satisfied are you with the following communication and dissemination tools?

- The quality of the Belmont Forum website as [*very high quality/very low quality*]
- The quality of the Belmont Forum BFgo as [*very high quality/very low quality*]
- The usefulness of the open access policy and requirements of the Belmont Forum as [*very useful/not useful at all*]
- The usefulness of the data-sharing tools of the Belmont Forum as [*very useful/not useful at all*]

Sustainability

15. [*Note to interviewer: Important question]* How **sustainable are the results achieved** (by your own project) so far?

- Are these outputs accessible and useable for others?
- What does the Forum do in order to encourage the continued use of results?

Conclusion

16. Is there **anything else** you would like to mention? (*also check comment box at the end of the survey*)

Appendix E Case studies

The case studies presented here are in-depth assessments of chosen thematic areas. For this evaluation, two case studies have been conducted, namely:

- A case study on North-South collaboration
- A case study on the complementarity and external coherence of the Forum

For both assessments, we have triangulated the data collected through desk research, interviews, survey, and bibliometric analysis.

A.12 A case study on North-South collaboration

The purpose of the first case study is to investigate the extent of North-South collaboration within the Forum. This should help answer the evaluation question concerning efficiency and organizational setup of the Forum, in particular the extent of inclusiveness.

A.12.1 *Design and methodology*

The case study relies on three sources of information and data:

- The survey of members and partners, which has been analysed for views on inclusiveness and also perspectives between members of the Global South and Global North.
- A number of in-depth interviews have also contributed to understanding the extent of inclusiveness.
- Bibliometric analysis of the Forum scientific publications to examine participation by researchers from the Global South.

A.12.2 *Participation of the Global South in the Belmont Forum*

A.12.2.1 Participation as members and partners

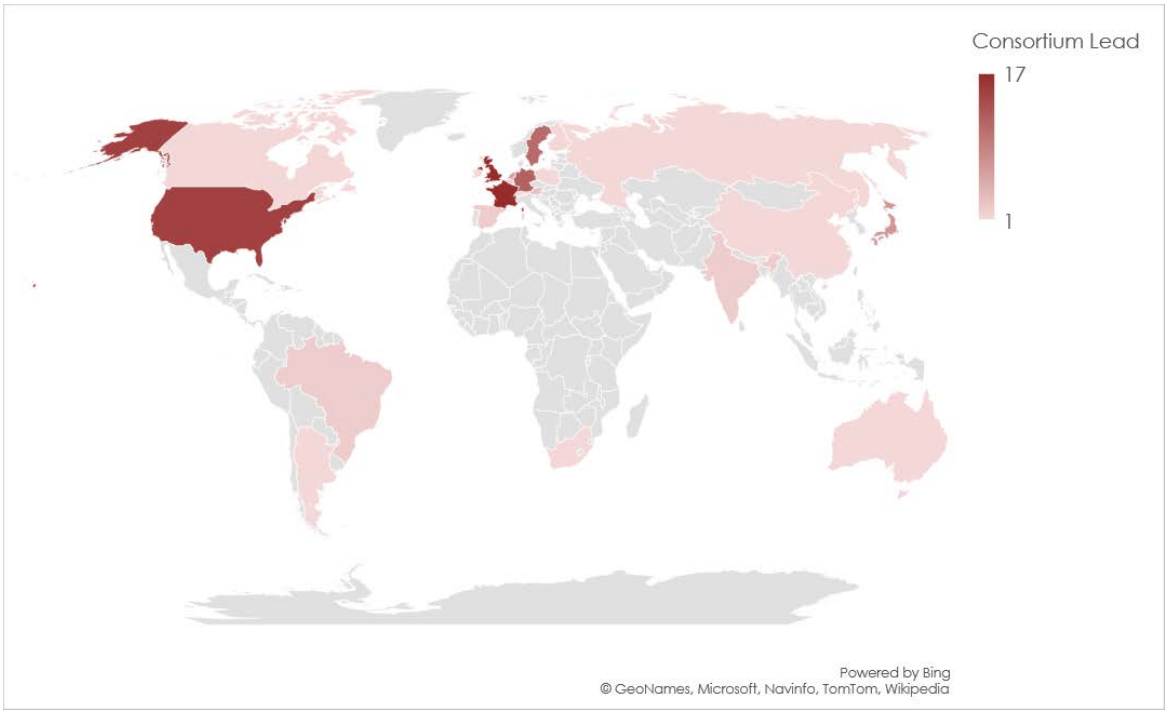
Of the 28 members of the Forum, 10 come from the Global South. The Global South members include the public research funding agencies from Argentina, Brazil[45], China, India, Ivory Coast, South Africa, Thailand and Turkey, as well as Inter-American Institute. From the beginning there has been substantial representation from the Global South with Brazil, India and South Africa as original members in 2012. Since then other members have joined from both North and South.

The funding of the CRAs is, nonetheless, strongly provided by members of the Global North, accounting for almost 90% of the total (according to the Secretariat).[46]

[45] For ease, we refer to Brazil although it is the state Sao Paulo Research Foundation which is a member, not a federal organisation.

[46] This is an estimate. Not all financial contributions to CRAs are recorded by the Secretariat and some members apparently would prefer that financial contributions are not disclosed.

Figure 1 Project consortium leads per country member (2012-2020)



Technopolis Group, 2020, based on data provided by the Secretariat (note that this does not include all information from ongoing CRAs)

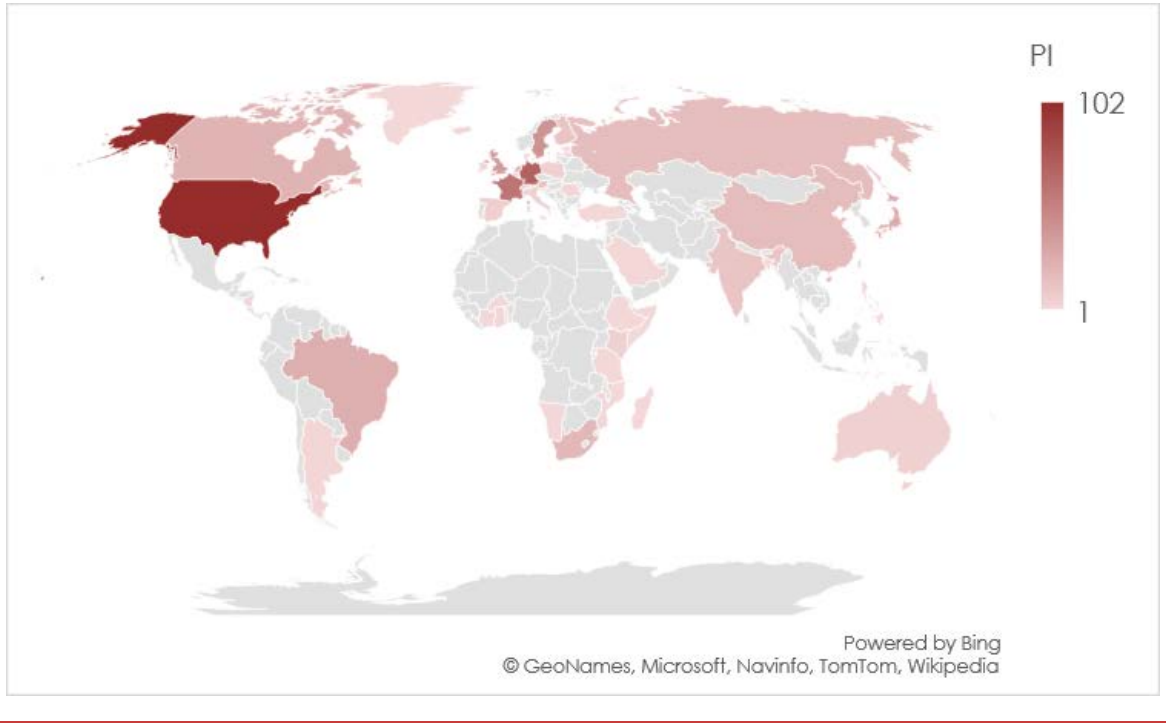
Members and partners of the Forum recognized this limited membership. In the survey of members and partners, two-thirds of the members felt that the geographical spread of the Forum's members was too limited and needed to be diversified. The findings from the interviews with members and partners indicate that the dissatisfaction is focused mainly on the limited participation from countries and researchers from the Global South (see also chapter 4.2.2.1 of the main report).

A.12.2.2 Participation as beneficiaries

Despite the high concentration of countries from the Global North, it is important to highlight that the Forum, nonetheless, did in fact manage to establish collaboration between the Global North and the Global South, more than many other funding instruments and initiatives did. The Forum has provided an opportunity for participation by researchers from and in the Global South in global environmental change research. We have not identified another funding mechanism which has provided this.

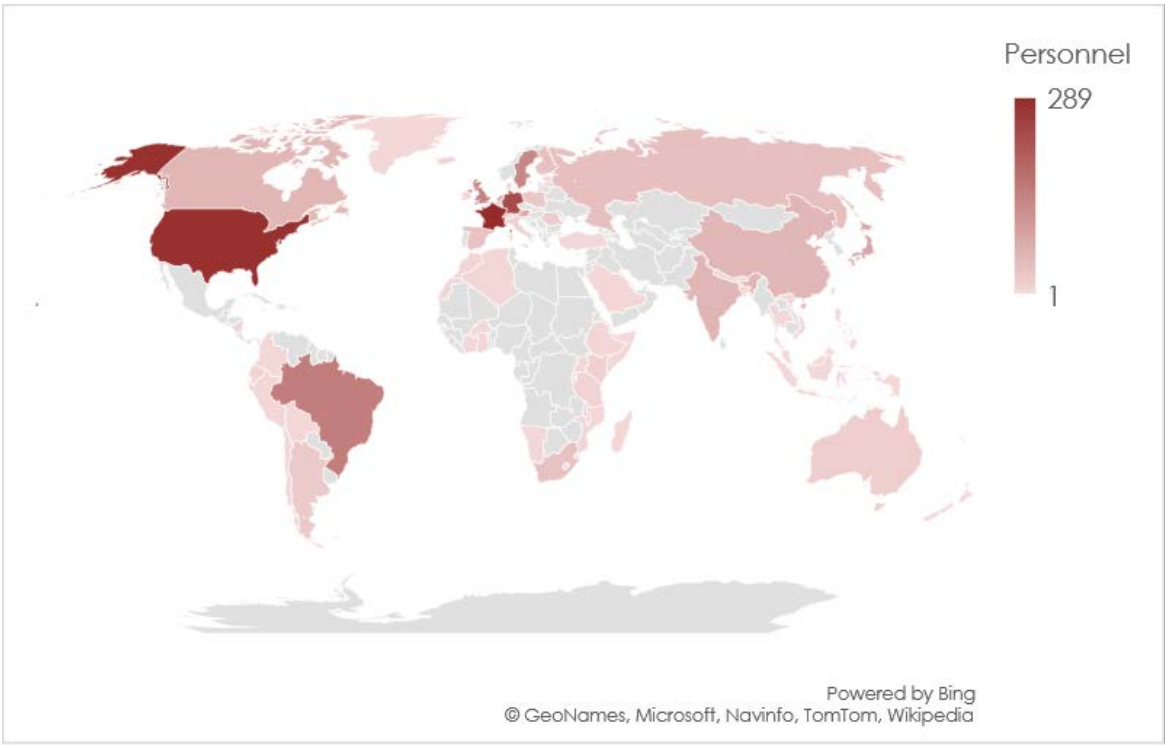
There is a broader participation of the Global South in projects, with various project partners coming from countries whose funding agencies are neither members nor partners of the Forum. This can be seen by comparing the distribution of total consortium leads, PIs and project personnel (see *Figure 1* to *Figure 3*). Consortium leads have been dominated by member countries of the Global North. The distribution of PIs shows however stronger representation from members of the Global South. Even more important, there are PIs from various countries in the Global South who are not members, particularly, though not exclusively, in Africa. This broader engagement of researchers from the Global South is even more prominent in the distribution of total personnel involved in CRA projects, with broader representation from countries in Latin America and South and Southeast Asia. This is compelling evidence that the Forum is providing additional opportunities for participation by researchers in the Global South.

Figure 2 Total project PIs per country (2012-2020)



Technopolis Group, 2020, based on data provided by the Secretariat (note that this does not include all information from ongoing CRAs)

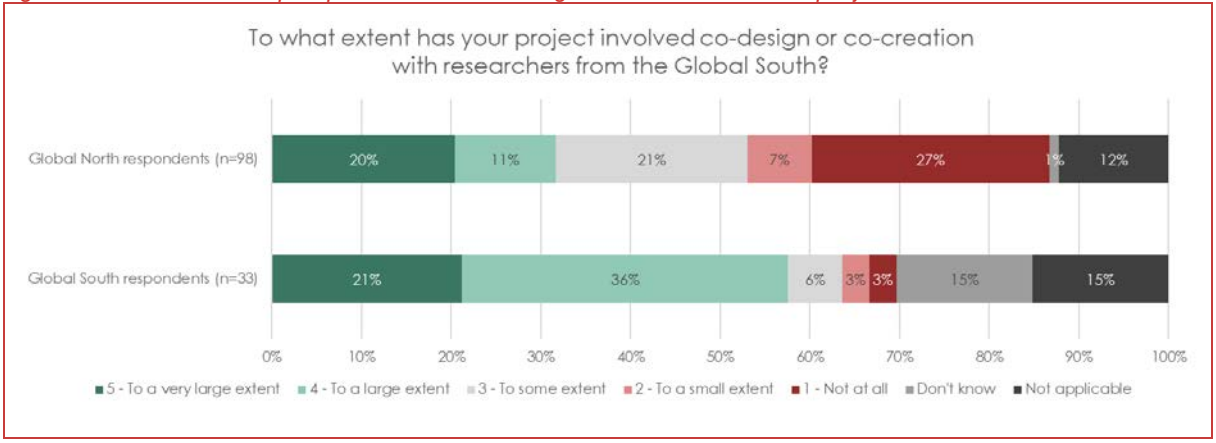
Figure 3 Total number of project personnel per country (2012-2020)



Technopolis Group, 2020, based on data provided by the Secretariat (note that this does not include all information from ongoing CRAs)

Furthermore, two-thirds of beneficiaries from the Global South who responded to the survey indicate that their project involved co-design or co-creation to a large extent or more (see *Figure 4*). Only 3% of total respondents from the Global South reported no co-design nor co-creation at all. The perspective of these researchers is arguably most relevant. For their part, about half of the researchers from the Global North indicated that partners from the Global South were involved in co-design or co-creation to at least some extent. About one fifth of these respondents from both regions judged that this co-design or co-creation with researchers from the Global South took place to a large extent. Although about a quarter of researchers from the Global North indicated that there was no co-creation or co-design with researchers from the Global South, this group could well include projects where that was less applicable. It is also not known whether the respondents from both regions were involved in the same projects, given that not all beneficiaries completed the survey. Nonetheless, the findings suggest that the Forum has succeeded in promoting co-design and co-creation on the part of the Global South. We have not identified another initiative with such data against which this conclusion can be benchmarked and indeed the absence of such an initiative with the same emphasis on co-design and co-creation substantiates our observation.

Figure 4 Beneficiaries' perspectives on co-design and co-creation in projects



Technopolis Group (2020)

A.12.2.3 Co-publication between North and South

The Forum may have contributed to increased joint publishing opportunities for participating researchers from the Global South. Almost half of Forum publications involve a combination of authors from the Global North and Global South. Only 3% of publications are authored exclusively by researchers from the Global South, as compared to 27% at a world level, or 24% for all the comparator funders analysed. There is though no evidence from the bibliometric analysis, however, of increased lead authorship for researchers from the Global South. 31% of Forum publications had either a lead or corresponding author from the South. This proportion compares similarly with what the same researchers achieved prior to the Forum and is actually somewhat lower than the wider group of global publications in similar thematic areas. However, this latter group might not be the most appropriate comparison as it includes non-international publications from large countries with stronger research activity and outputs, such as China, India and others.

Of greater significance is the result of the bibliometric analysis showing that North-South international co-publication rates exceed those of almost all funding groups taken as points of comparison. This is also the case with international co-publication rates (ICR), while the North South ICR is calculated including at least one author coming from the Global South. For the

Forum, the North-South ICR is 41%. The average of comparator groups is 18%. The only other comparator groups with rates above 40% are the Chinese and Brazilian funding agencies included in the analysis. It seems unlikely that such a level would be seen among the participating researchers from the Global South in the absence of the funding and collaboration mechanisms provided by the Forum. This is supported by the comparison with non-Belmont Forum publications by beneficiaries of the Forum projects, which is 35%. The North-South ICR for Forum beneficiaries prior to their grants was 27%. The importance of these results is strengthened by the overall positive results in terms of citation scores for Forum publications (as detailed in the bibliometric analysis).

A.12.3 *Constraints and challenges*

The Forum has thus achieved substantial results in terms of inclusion amongst its current members from the Global South. Broadening inclusion of the Global South requires more members and a means, or new funding models, to support their participation, particularly for researchers in low-income countries. A principal constraint is that funding agencies from the Global South generally have far fewer resources than those in the Global North, especially in the case of low-income countries. The model of the Forum is based on members and partners mobilising and contributing their respective resources. This model entails challenges to promote greater inclusion of researchers from the Global South.

Some specific constraints mentioned in interviews include cultural differences. It was reported that participants from the Global South may find it difficult to work at the same pace as Northern members and partners in the CRA scoping process. This suggests that the Forum works according to an organizational culture that provides challenges for some members from the Global South. The interviews also indicated that for some existing and potential members from the Global South, the increased partnership and engagement with other stakeholders, such as civil society, can make participation of the members of the Global South more problematic.

The interviews also produced the observation that the Global South is not homogenous. The most gains in participation have been achieved among Asian members, as compared to Africa or Latin America. This suggests that in looking for possibilities to enhance participation by the Global South, a regional-specific approach may be needed.

A.12.4 *Recommendations*

Expanding the membership from the Global South would help enhance its participation, making the research more international and possibly relevant for understanding and adapting to global environmental change affecting these countries. It may not be realistic though to expect substantially matching contributions from members of the Global South and therefore a mechanism and funding needs to be found to support their participation. Suggestions made include sponsorships for researchers from the Global South. This and other funding mechanisms to promote such participation might involve engagement with other organizations, such as development assistance organizations or philanthropic foundations.

The Forum could also review the CRA scoping process to identify ways to encourage greater participation of Global South. According to the Secretariat, some efforts have been made, including working with partners, such as SIDA, to target funds for lower and middle income country participation in the Oceans, CEH and Pathways CRAs.

A.13 A case study on the complementarity and external coherence of the Forum

The purpose of the second case study on complementarity is to assess the extent to which the activities of the Forum are coherent with other initiatives in the realm of global environmental change. This should help answer the evaluation question concerning external coherence of the Forum. Given the range of other funding or coordination mechanisms operated by many Forum members, for instance Horizon 2020 or Future Earth, there appears to be a possible risk of overlap, which might appear to be more likely from the perspective of an external observer.

A.13.1 *Design and methodology*

The case study relies on three sources of information and data:

- Desk study to map the landscape of relevant initiatives and assess the positioning of the Forum.
- The survey of members and partners, which has been analysed for views on added value and relationship to other initiatives.
- A number of in-depth interviews, which have also contributed assessing the possible gaps and overlaps between the Forum and other initiatives.

A.13.2 *Landscape of international initiatives*

Relevant international initiatives in the landscape within which the Forum operates are listed in Table 1. These initiatives are relatively diverse in terms of the type and scope of their membership. Some, such as the International Science Council (ISC), are organisations that bring together international scientific unions and associations. The ISC[47] was created in 2018 as the result of a merger between the International Council for Science (ICSU) and the International Social Science Council (ISSC). Members are essentially national academies of science. The ISC has also become partner of the Forum.

Future Earth is an initiative that grew out of three previous international research programmes: the International Geosphere-Biosphere Programme (IGBP), the International Human Dimensions Programme on Global Environmental Change (IHDP), and DIVERSITAS, which were all under the umbrella of the Earth System Science Partnership (ESSP), an initiative of the former International Council for Science (ICSU). The Forum is a member of the Governing Council of Future Earth.

Various funding organisations are not included here given their relatively incomparable governance structure. This includes "quasi-national" organisations such as the International Development Research Centre (IDRC), established by Canada, and the Institute for Global Environmental Strategies (IGES), established by Japan. Such organisations are active in the same landscape and have a somewhat international character. They differ though considerably from international initiatives and partnerships, which are the focus here, in that the former are primarily funded by one government. The Forum has had interactions with both of these organizations about possible partnership in CRAs.

[47] https://council.science/

A.13.3 *Analysis of key initiatives in relation to Belmont Forum*

The **mission** of the Forum could be formulated as, "*creating the global collaboratory for research in support of environmentally sustainable development*"[48] More specifically, the Forum supports research that takes account of coupled natural, social, and economic systems, with a wide range of objectives as listed in the Belmont Challenge White Paper.

The mission and objectives of the other initiatives are listed in Table . In comparison, the Forum stands out as the primary initiative to support research on global environmental change at an international level. The ISC promotes research and aims to be the voice of science, but does not have mechanisms for financing research. Nor does the ISC have a specific, thematic mandate. The main objectives of the Global Research Council are to share data and promote cooperation among science funding agencies. The WCRP promotes international science coordination and partnerships in the area of climate science, and is connected to the UN system, including the WMO, UNESCO and the UNFCCC.

The Forum has a mission to support **transdisciplinary** research, which is distinct from other initiatives listed above. There are various initiatives that promote scientific collaboration, but the Forum is the only one that has funding instruments and that is focused on global environmental change.

The mission of Future Earth – "to accelerate transformations to global sustainability through research and innovation" – does seem to overlap with that of the Forum. Future Earth operates, however, more as a platform for researchers than as means to support and fund research. Future Earth also has objectives of building and mobilising networks and shaping the global narrative, moving to bridge the science-policy interface. Given these differences, it is understandable and logical that the Forum has been instructed through its governance to collaborate with Future Earth. This collaboration takes various forms, one which is the Sustainability Research and Innovation Congress, an annual gathering for transdisciplinary researchers, implementers, funders, and coordinators to support networking, action, training, collaboration, and broader engagement.

The objectives of certain regional initiatives are also similar to those of the Forum. The Joint Programming Initiatives (JPIs) of the European Commission aim to pool national research efforts in order to make better use of respective national research and development resources to tackle common challenges more effectively. The Forum has partnered with specific JPIs in various CRAs (including FWE-Nexus, Oceans, Climate Predictability, and Food Security). The JPIs are however, limited to the EU. The EC's ERA-NETs under Horizon 2020 are included in the table partly for historical purposes. These instruments, including BiodivERsA, have shifted from funding networks to one of "topping-up" funding of single joint-calls for transnational research and innovation (termed "co-funds"), making them less relevant for comparison to the Forum. There has been collaboration with BiodivERsA in the past in some CRAs (including T2S, Biodiversity II, FWE-Nexus).

In terms of research topics, the Forum concentrates exclusively on international transdisciplinary research for understanding, mitigating and adapting to global environmental change. The Forum is thus the only initiative that leads to joint funding of research at a trans-continental level, on environmental change, and with a transdisciplinary focus. There does not appear to be a clear alternative mechanism by which this could be achieved. Two-thirds of members

[48] Belmont Forum Terms of Reference (2015)

and beneficiaries who responded to the survey indicate that a reason for participation in the Forum is to realise international cooperation that would otherwise not be realised.

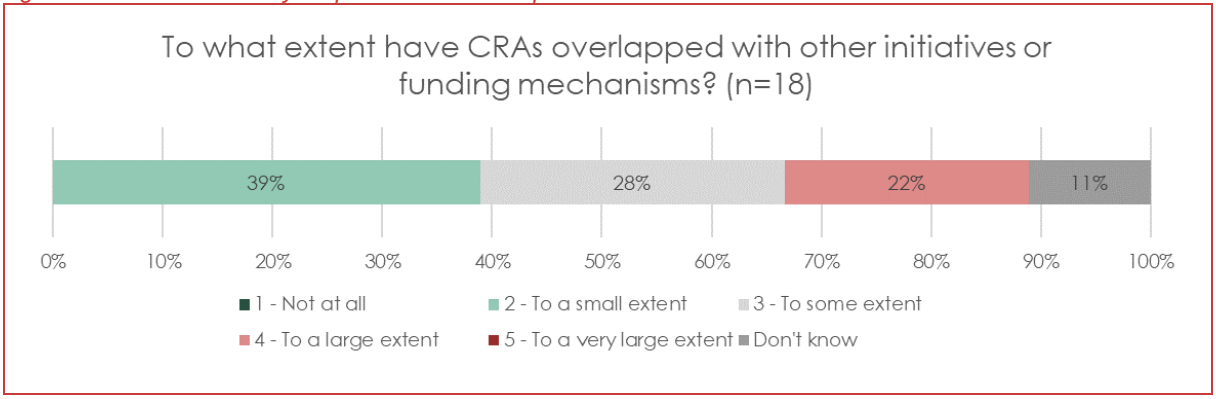
A.13.4 Overlaps and complementarities with other initiatives.

Although the Forum is unique in its mission, scope and membership, various overlaps with other initiatives have been reported by respondents to the survey (see *Figure 1*). These refer to specific CRAs, though it is not clear from the survey results which of the CRAs this concerns. In general, it seems reasonable and to be expected that there is some overlap (overlap to a small extent or to some extent) between CRAs and some other research and funding initiatives. There can be planning considerations or programmatic issues that inhibit another initiative from joining a CRA.

Partial overlap can occur on some dimensions but not others. For example, other initiatives are funding research on global environmental change, but this might not be as transdisciplinary as that funded by the Forum. An initiative such as Future Earth is promoting transdisciplinary research but does not itself fund calls for proposals. Partial overlap can occur in terms of scope. The Forum currently has an international membership with global scope, yet not as much representation from the Global South. Individual CRAs could overlap somewhat with specific regional initiatives, such as JPIs, though there has been partnering with some CRAs, as mentioned above.

Where members and partners report overlap "to a large extent", this might be interpreted as suggesting that there may be duplication of efforts in some areas. This includes a quarter of the members who responded to the survey and had an opinion about this. This group of members includes major funding agencies and is spread across continents. It includes one agency which was a founding member as well as more recent members. Thus, this perspective, while limited to a quarter of members, is not confined to a specific region or other clear common characteristic among these respondents. At the same time, the nature of perceived overlap is not fully understood. The Forum is designed to connect with national and regional initiatives, and to leverage these with funding for international collaboration. It would be insightful to assess further whether members are of the opinion that this overlap should be reduced, or improved in some other ways.

Figure 1 Member survey responses on overlap with other initiatives



Technopolis Group (2020)

Many individuals who were interviewed indicated that coordination with other initiatives should be improved, but there are few specific suggestions about how to do that. Clearly increased coordination also requires resources for the Secretariat. It should be noted that the Secretariat

has engaged with external coordination efforts for some time, and has suggested the creation of specific coordination offices for each CRA.

Overlap seems most apparent with other funding mechanisms or organisations that fund a similar kind of transdisciplinary environmental change research at an international level, such as specialised agencies, including IDRC, and specific philanthropic organisations. One way to reduce this overlap would be to expand membership of the Forum, attracting such organisations as partners, perhaps to fund specific CRAs. This evaluation has not explored whether other organisations are interested in such an association. Their own independence may well be seen as part of their relative advantage. Membership is only option. There can be other forms of partnering, particularly on specific CRAs. Indeed, the Forum has explored such opportunities and there are some examples, such as the Pathways CRA with which GEO has partnered.

According to interview responses, the most important areas in which **complementarity** could be enhanced are with Global Research Council (GRC), Future Earth, and connections with other types of funding agencies, including organisations such as IDRC, development aid agencies and philanthropic foundations. A substantial number of interviewees indicate some overlap between the GRC and the Forum and also indicate that there should be more clarity on their relative roles. This seems curious in some ways as the GRC does not coordinate calls for funding and was even described in one interview as being too political. It seems that it may be more the case that there is a need for clearer communication of the relative roles of the range of initiatives in general and how they complement each other. Indeed, these observations are noteworthy given that most interviewees would be among those who would be expected to understand best the landscape of initiatives within which the Forum operates.

The **relationship with Future Earth**, which is complex, seems to be a key factor in complementarity and external coherence. At the most basic level, the two initiatives complement each other in terms of their function and activities. To an external observer, the question is why these two initiatives exist separately? There are clear reasons, not only historical circumstances, for this, including the separation of funding agencies from other actors. Interviewees mentioned the need for the relationship between the Forum and Future Earth to be more clearly articulated. It seems compelling thus that this relationship could be more strongly and clearly articulated and communicated to all stakeholders, particularly those who are only partly engaged in these communities (for example, policy advisors and other users of research outputs) as well as researchers who may be (potential) beneficiaries but are not involved with coordination at the international level at which these organizations operate.

Other initiatives mentioned by interviewees include the International Science Council (ISC), which brings together a wider range of scientific academies and organizations, among which are some funding organisations. There was a suggestion among interviewees to develop more of a global forum of funders of global environmental change research.

As with addressing overlaps, enhancing complementarity requires resources for the Secretariat, as well as instructions and oversight from the members and partners. In practical terms, enhancing complementarity would essentially be at the level of the CRAs. This suggests finding ways to improve the scoping process for CRAs to enhance complementarity. This could involve additional analysis and justification of how a proposed CRA both does not overlap with other initiatives and in which way it complements existing initiatives. The Secretariat reports some improvements from the new virtual scoping process. The CRA proposal template requires information about how the theme connects with existing programs and how to leverage any existing activities. This evaluation has not determined though whether these requirements are subjected to sufficient scrutiny and analysis.

A.13.5 *Recommendations*

In summary, reducing real or perceived overlaps and enhancing complementarity with other initiatives could be achieved at two levels:

- **External engagement and communication** by the Secretariat: This clearly requires resources for the Secretariat. While there seems to be relatively little overlap with other initiatives in general, a priority area for enhancing and communicating complementarity is the relationship with **Future Earth**, among both the international scientific community and broader, associated stakeholders.
- **Refining the scoping process for CRAs**: Possible areas to explore are the examination of the selection criteria to ensure a stronger and more apparent complementarity to other initiatives.
- **Synthesis of complementarity**: An analysis and synthesis of past CRA outcomes in relation to other programs and initiatives could inform potential themes going forward, enhancing external complementarity.

Table 1 Overview of related initiatives

Initiative name	Description	Members/partners	Mission and objectives	Types of funding activities	Research topics	Geographical scope
Future Earth	Global network of researchers collaborating on sustainability research, funding knowledge action networks and global research projects	No members; partners include international research funders, national funders, national institutes and universities	Global networks of researchers collaborating on sustainability research. Mission: to accelerate transformations to global sustainability through research and innovation. Objectives: - Facilitate research and innovation - Build and mobilize networks - Shape the Global Narrative.	• Knowledge-Action Networks • Global Research Projects Research projects generate scientific knowledge. Networks link policy, business and civic leaders with researchers (building partnerships) for better translation of knowledge to practice.	Emergent risks and extreme events; finance and economics; health; natural assets; ocean; systems of sustainable consumption and production; urban; water-energy-food Nexus.	Broad, international: 5 global hubs, 5 regional centres and offices, 20 global research projects.
International Science Council (ISC)	Non-governmental organisation with 40 international scientific Unions and Associations and over 140 national and regional scientific organisations. One of their initiatives is The Global Forum of Funders, which aims to set up a common call for a decade of global sustainability funding action.	40 international scientific Unions and Associations and over 140 national and regional scientific organizations including Academies and Research Councils.	The mission is to be the global voice for science; a trusted voice that speaks for the value of all science by: 1. Promoting international research and scholarship on key global challenges; 2. Increasing evidence-informed understanding and decision-making at all levels of public policy, discourse and action; 3. Promoting the continued and equal advancement of scientific rigour, creativity and relevance in all parts of the world; 4. Protecting scientific freedom and advocating principles for the responsible practice of science	Three principal areas of work. • Science-for-policy to stimulate and support international scientific research and scholarship, and to communicate science that is relevant to international policy issues. • Policy-for-science to promote developments that enable science to contribute more effectively to major issues in the	4 research domains: The 2030 Agenda for Sustainable development. The digital revolution. Science in policy and public disclosure. Evolution of science and science systems.	Broad, international, offices in Africa, Asia and the Pacific, and Latin America and the Caribbean

				international public domain. • Scientific freedom and responsibility to defend the free and responsible practice of science.		
Global Research Council	Organisation comprised of the heads of science and engineering funding agencies from around the world, with the objective of fostering multilateral research and collaboration across continents	Heads of science and engineering funding agencies from around the world	The Global Research Council aims to promote the sharing of data and best practices for high-quality collaboration among funding agencies worldwide. Goals: To improve communication and cooperation among funding agencies; To promote the sharing of data and best practices for high-quality research cooperation; To provide a forum for regular meetings of the Heads of Research Councils; To respond to opportunities and to address issues of common concern in the support of research and education; To be a resource for those institutions wishing to build a world-class research landscape; To explore mechanisms that support the global science enterprise and the worldwide research community.	No funding calls.	1. Building Research and Education Capacity 2. Capacity building and connectivity and granting agencies worldwide 3. Interplay between fundamental research and innovation 4. Scientific breakthroughs 5. Interdisciplinarity 6. Open access 7. Gender equality in research 8. Research integrity	International, broad.
Asian Pacific Network for Global Change Research	Network of member country governments that promotes global change research in the region, supporting research and science-based response strategies in the field of biodiversity, climate	National research funding organisations from 12 member countries.	The Asia-Pacific Network for Global Change Research (APN) is a network of member country governments that promotes global change research in the region, increases developing country involvement in that research, and strengthens interactions between the science community and policy makers.	support for research and science-based response strategies and measures, effective linkages between science and policy, and scientific capacity development	9. Biodiversity and ecosystem services 10. Climate adaptation framework 11. Low carbon initiatives	Pan-Asian

	adaptation and low carbon initiatives		The mission of APN is to enable investigations of changes in the Earth's life support systems and their implications for sustainable development in the Asia-Pacific region through support for research and science-based response strategies and measures, effective linkages between science and policy, and scientific capacity development. APN, therefore, supports investigations that will: Identify, explain, project and predict changes in the context of both natural and anthropogenic forcing; Assess potential regional and global vulnerability of natural and human systems; and Contribute, from the science perspective, to the development of policy options for appropriate responses to global change and sustainable development.			
European Strategy Forum on Research Infrastructures (ESFRI)	Self-regulated body consisting of delegates of Member States and the European Commission. Focus on streamlining research infrastructures for a large number of research areas, including energy, food and environment	Representative of each EU Member State+EC	Self-regulated body consisting of delegates of the Member States (science-policy officials) and EC. Mission: to support a coherent and strategy-led approach to policymaking on research infrastructures in Europe, and to facilitate multilateral initiatives leading to the better use and development of research infrastructures, at EU and international level. Objectives: • to support a coherent and strategy-led approach to policy making on research infrastructures in Europe; • to facilitate multilateral initiatives leading to a better use	SFRI operates at the forefront of European and global science policy and contributes to its development translating political objectives into concrete advice for RI in Europe. No funding activities for researchers as such., focus is on streamlining research infrastructures. Several RI initiatives supported: https://www.esfri.eu/ri-initiatives	Energy; health and food; environment; social and cultural innovation, physical sciences and engineering; data, computing and digital research.	EU

			and development of research infrastructures acting as an incubator for pan-European and global research infrastructures; • to establish a European Roadmap for research infrastructures (new and major upgrades, pan-European interest) for the coming 10-20 years, stimulate the implementation of these facilities, and update the Roadmap as the need arises; • to ensure the follow-up of implementation of already ongoing ESFRI projects after a comprehensive assessment, as well as the prioritisation of the infrastructure projects listed in the ESFRI Roadmap.			
European Joint Programming Initiatives (JPI)	Strategic Research Agenda of European Member States with specific thematic focus: 12. JPI Climate: Connecting Climate Knowledge for Europe 13. JPI Oceans: Healthy and Productive Seas and Oceans		The aim of the joint programming process is to pool national research efforts in order to make better use of Europe's research and development resources and tackle common European challenges more effectively. JPIs are developed in a structured and strategic process where EU countries agree on a voluntary basis on common visions and Strategic Research Agendas (SRA) to address major societal challenges.			
European Commission ERA-Net Joint funding schemes	Network programming and funding research across EU (countries and territories) including funding. Specific schemes of relevance include: • BioDiversa: biodiversity and ecosystem services	39 (European+) national and regional funding organisations	BiodivERsA - the network programming and funding research on biodiversity and ecosystem services across European countries and territories BiodivERsA is a network of 39 agencies and ministries from 25 European countries and associated countries programming and funding pan-European research on biodiversity	The network's activities are organized under 11 work packages (WPs): WP1: Reinforcing the network and the links to other relevant research network and funders (WP leader : FCT, Portugal) WP2: Integrating research programmes and addressing	'biodiversity and ecosystem services and their valuation' (9.6M€), 'biodiversity scenarios, identifying tipping points and improving resilience' (8.8M€), 'invasive species and biological invasions' (8,9M€) and 'promoting synergies	EU + some associated countries

	FACCE SURPLUS: Sustainable and resilient agriculture for food and non-food systems		and ecosystem services on a competitive basis. It is an ERA-NET Co-fund, funded under the EU's Horizon 2020 Framework Programme for Research and Innovation. BiodivERsA's objectives are to: Reinforce the network and its capacity to join-up research on biodiversity and ecosystem services in Europe and in overseas regions and territories Develop a strategic, multi-annual vision of the network's priorities based on ambitious mapping and foresight activities, linking existing international and European agendas with national and institutional priorities. Work towards the alignment of research programmes on biodiversity and ecosystem services Implement joint calls for research proposals and instate a recurrent, well-identified funding scheme for transnational biodiversity research projects, allowing to better integrate research on biodiversity and ecosystem services in Europe Promote an effective interfacing of science with society and policy during the whole research process, through the co-design and joint implementation of research programmes and research projects it funds.	issues with and within Overseas territories / Outermost regions (WP leader : Guadeloupe Region, France) WP3: From mapping and foresight activities to the development of a common strategic roadmap (WP leader : NWO, The Netherlands) WP4 to 8: Implementation of the 2015 joint call (WP leader : ANR, France) WP9: Implementing other joint activities (WP leader : Formas, Sweden) WP10: Science-society / science-policy / science-business interactions and communication (WP leader :BelSPO, Belgium) WP11: Coordination and management (WP leader :FRB, France)	and reducing trade-offs between food supply, biodiversity and ecosystem services' (10.2M€ - call jointly launched with FACCE-JPI).	
Institutionalised European Partnerships (article 187, 185)	Multiannual programme jointly implemented by EU Member States and Associated Countries, with a focus on European Metrology. Long term public-					

	public partnership with the ambition to achieve scientific, managerial and financial integration amongst national research programmes in a given field.					
Mission-Innovation	25 members on 5 continents, working to stimulate innovation with the objective to make clean energy widely affordable. Implementation of Innovation Challenges through bilateral or multilateral activities, including joint R&D calls	24 countries + EC. Countries represented by Ministers of energy-related organisations within government.	25 members on 5 continents are working to stimulate innovation with the objective to make clean energy more widely affordable	Innovation Challenges (ICs) are implemented through bilateral or multilateral activities include joint RD&D calls, demonstration projects, student and researcher exchanges and coordinated funding calls.	Innovation Challenges: Smart Grids Off-grid access to electricity Carbon Capture Sustainable Biofuels Converting Sunlight Clean Energy Materials Affordable Heating and Cooling Buildings Renewable and Clean Hydrogen	International, broad.

Appendix F Technical Report: bibliometrics and altmetrics

To explore tangible scientific outputs of the Forum's activities, we have included a bibliometric and altmetric analysis into this evaluation. The analysis was performed by Science-Metrix. Their final report can be found on the next page

A.14 Contents of the bibliometrics and altmetrics report

Tables

Figures

Acronyms and abbreviations

AMI: Altmetrics mentions index

ARC: Average of relative citations

BF: Belmont Forum

BMBF: German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung)

CDC: Citation distribution chart

CDI: Citation distribution index

CRA: Collaborative Research Actions

Ctries: Countries

EC: European Commission

ERC: European Research Council

EU: European Union

FAO: Food and Agriculture Organization of the United Nations

ICR: International co-publication rate

IGO: Inter-governmental organization

II: Interdisciplinarity index

IPBES: Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services

ISR: Inter-sectoral co-publication rate

IUCN: International Union for Conservation of Nature

$HAP_{10\%}$: Share of highly altmetrics-mentioned publications

$HAP_{1\%}$: Share of highly altmetrics-mentioned publications

$HCP_{10\%}$: Share of highly cited publications

$HCP_{1\%}$: Share of highly cited publications

$HIP_{10\%}$: Share of highly interdisciplinary publications

$HMP_{10\%}$: Share of highly multidisciplinary publications

MI: Multidisciplinarity index

NGO: Non-governmental organisation

NERC: UK Natural Environment Research Council

NSF: US National Science Foundation

NSFC: National Natural Science Foundation of China

OA: Open access

Pol.: Policy

Pubs: Publications

UNEP: UN Environmental Programme

WCS: Weighted CiteScore

WMO: World Meteorological Organization

Introduction to the bibliometric and altmetric analyses

The Belmont Forum (the Forum) aims to bring together academics with stakeholders and end users of research in a co-production approach. It aims to foster high levels of North-South collaboration in doing so. The projects supported revolve around the environmental sciences but extend to many other disciplines, spanning a range from advanced mathematical climate modelling to the arts and humanities. Finally, the program aims to support the production of research outcomes with high societal and local relevance as much as the more traditional outputs of scientific excellence.

Recap on the overall evaluation approach

The inception report submitted in modified form on January 30, 2020, provides the basic coordinates for the conduct of the current External evaluation of the Forum. The following two subsections are reproduced from the inception report, so that the broader evaluation context be better taken into account when interpreting the findings provided in this report.

Rationale and objectives of the evaluation

The Forum has commissioned an external evaluation of the Forum. Based on the Request for Proposals, the main objective of this evaluation is: *"to evaluate its progress towards meeting the Belmont Challenge and the efficacy of the organisation to continue to reach its goals and fulfil its mission"*. More specifically, the evaluation will:

- Demonstrate the Forum members the **benefits and added value from investment** in the Forum;
- Help the Forum to demonstrate the **added value of the transdisciplinary, transnational approach** to Collaborative Research Action (CRA) partners (both member and non-member organisations);
- Evaluate the **flexible partnership methods** of the CRA process; and
- Validate the **benefit of transdisciplinary approaches for informing decision making**, policy, and practice.
- The evaluation will also provide recommendations for the future of the Forum.

Scope of the evaluation

The external evaluation of the Forum will cover the period from 2009 (when it was established) until mid-2019. The evaluation will consist of two main components:

- **Impact evaluation:** an assessment of the delivery of the CRAs, including their scientific, policy, and other impacts and the added value of transnational and transdisciplinary collaborations. In this context it should be noted that the first projects have only finished very recently, which means that the assessment of impact will be somewhat limited (as it may not be visible yet).
- **Organisational evaluation:** an exploration of the effectiveness of the Forum (including the funding mechanisms, its governance and management (Plenary, Steering Committee and Secretariat) efficiency of procedures, transparency and inclusiveness, etc.).

In order to structure this evaluation, we defined six broad evaluation criteria, namely:

1. **Relevance:** the extent to which the mission, objectives and activities of the Forum are in line with needs and problems,

2. **Effectiveness and impact:** the actual effects, results, and impacts achieved by the Forum and its funded projects so far including the extent to which the Forum reaches out to stakeholders and society.

3. **Efficiency and organisational set-up:** the extent to which the results are achieved at a reasonable cost (including resources), the extent to which the governance, management and organisation of the Forum are adequate and the extent to which the Forum's processes and procedures are smooth and time efficient.

4. **Coherence:** the extent to which the CRAs and selected projects complement each other, and the extent to which the Forum as a whole complements other (global) initiatives with similar objectives.

5. **Added values:** the extent to which the results and impact observed could (would) not have been achieved in the absence of the Forum.

6. **Sustainability:** the extent to which the Forum itself, but also results of funded projects are sustainable, even if the Forum were at some point cease to exist.

The use of bibliometrics for evaluating a transdisciplinary, multilateral organisation: benefits and challenges

In this project, a major challenge was the development of quantitative evaluation indicators for participatory, collaborative and local research outcomes. Bibliometrics approaches, conducted on traditional journal-based research outputs, benefit from the availability of a small number of central databases with comprehensive records of scientific publications. There is no centralised documentation mechanism in place for participatory and local research outcomes, greatly complicating data collection and benchmarking. Emerging altmetric methods tread a middle ground, capturing online interactions towards traditional scientific publications.

As the findings presented below will show, a sizable portion of the Forum-supported projects have produced journal-based publications as part of their outcomes, alongside other traditional academic outputs (including conference presentations and posters, or graduate student theses). Retrospectively, bibliometrics clearly appeared as an essential component for harvesting the Forum-supported research outcomes.

Societal outcomes of research are most often appraised using qualitative approaches, including case studies, interviews, and expert panels. Surveys are also commonly used and may just as often be of a qualitative as of a quantitative character. Technopolis is using some of these approaches to try and capture a part of the societal outcomes originating in the Forum-supported projects. Science-Metrix' contribution to the appraisal of societal outcomes will be at once broader in scope but of more limited depth.

First, Science-Metrix was able to retrieve observations on social media, wiki, journalistic and policy uptake of the Forum-supported journal-based outputs, using altmetric approaches. Altmetrics are perceived as a highly promising toolbox to track at least portions of the societal outcomes of research.[49] Nevertheless, altmetrics can offer only a partial picture of these outcomes, and not

[49] European Commission Expert Group of Altmetrics. (2017). *Next-generation metrics: Responsible metrics and evaluation for open science*. Brussels.

all metrics retrieved are directly relevant to this goal.[50] The findings from these analyses will be presented with relevant notes for interpretation and clear reminders of limitations. As will be seen below, altmetrics activity around the Forum's journal publications was sustained and offered highly interesting insights into some of the potential societal outcomes to originate in supported projects.

Secondly, Science-Metrix collected web citation, hyperlink and usage statistics on non-journal outputs supported by the Forum. Outputs such as online videos, blog posts, web pages, webinars, policy reports and journalistic pieces were expected receive some attention from society-side users and stakeholders, attention that can be captured in online links, downloads and/or views. With the URLs of citing websites available, it was hoped that citations of special interest, originating from local communities, NGOs, partners, and the like, could be identified. This portion of the analysis retrieved comparatively low numbers of observations, and for the most part did not help identify cases of in-depth participatory or local uptake. It is unclear if altmetrics methods are not yet appropriate for use with non-journal outputs, or whether this is indicative of restricted outcomes here. Triangulation with future findings from the Technopolis-lead surveys and interviews may help to reduce uncertainty surrounding this issue.

Comparatives strategies and analytical periods

Access to baseline and benchmark data sets greatly increases the interpretative value of performance measurements performed as part of programme evaluations. Science-Metrix has examined the structure, features, and data available on programmes that could act as comparators to the Forum's CRAs for the purpose of the current evaluation (see a detailed description of the process in section 0).

Science-Metrix reached the conclusion that the BiodivERsA programme, funded through the European Commission's Framework Programmes (i.e., FP6, FP7, and H2020), would act as the primary comparator in the evaluation, for the following reasons:

- the duration of its awards;
- the monetary value of its awards;
- the use of a similar joint call structure that combines funders from multiple countries;
- the typical composition of supported teams (i.e., interdisciplinary and international); and
- the thematic focus of its awards.

Additional comparator groups were also identified with other research strategies, including the constitution of a thematic publication set revolving around publications in the key subfields where articles supported by the Forum are located; and by examining concurrent non-Forum as well as prior publication sets by investigators supported by the Forum.

Primary comparator using bibliometric and altmetric indicators on peer-reviewed publications: In recommending the use of the BiodivERsA programme as a primary comparator for the Forum's CRAs, Science-Metrix notes the availability of an in-depth review of outputs from the 2008–2009 joint call of that programme, published in 2015. This review contains comprehensive lists of project outputs, including publications and other forms of research by-products, which were, as for the Forum, desirable given the call's focus on linking scientific advancement to policy and practice.

[50] Robinson-Garcia, N., Costas, R., Isett, K., Melkers, J., & Hicks, D. (2017). The unbearable emptiness of tweeting—About journal articles. *PLOS ONE*, *12*(8), p. e0183551. doi:10.1371/journal.pone.0183551; Tahamtan, I., & Lutz Bornmann, ; (2020). Altmetrics and societal impact measurements: Match or mismatch? A literature review. *El profesional de la información*, *29*(1), p. e290102. doi:10.3145/epi.2020.ene.02.

Although the reference period for the review of the BiodivERsA call (i.e., 2008–2009) is prior to any of the Forum's CRA calls, it should be kept in mind that many of the standard bibliometric indicators computed by Science-Metrix are normalised by year and field of science, with the world as a reference.

Primary comparator using custom altmetrics on non-article outputs: For the specific case studies to be performed using custom altmetrics on non-article outputs, the comparison of the 2008–2009 call of BiodivERsA to the 2012 CRAs will be ideal since information on these types of outputs will, in both cases, have been gathered roughly seven to eight years following the respective call (i.e., in 2015 for BiodivERsA and in 2020 for the CRA). The 2015 BiodivERsA review, however, has appraised non-articles outputs in a narrative mode rather than a quantitative one. Data from the review can therefore only be used as a reference in the qualitative portion of the analysis. BiodivERsA project websites were nonetheless examined for information on project non-journal outputs, but multiple websites had closed since the completion of the research projects. Comparisons will only be made with caution in this subset of analyses.

Longitudinal analyses and within-group comparisons: Some comparisons aimed to measure the extent to which plans for transdisciplinarity and collaboration, as they appear in grant applications, have materialised in peer-reviewed scientific outputs. In those cases, publications supported by the Forum's investigators were retrieved for the five years period prior to the start of their award. Comparison were conducted on achievements during and prior to the award period.

Comparisons between pre- and award periods do not control for broader trends. For example, an increase in the international co-publication rate of the supported researchers could have occurred in the absence of the Forum since this is generally a clear pattern at world level. Also, it is possible that an effect is not observed because the Forum supported individuals who were already exhibiting a strong propensity towards international cooperation.

To help control for such confounding factors, a comparator group was also elaborated from parallel publications by the Forum investigators published during the award period but with no mention of the Forum funding in their acknowledgements. This control group's differences in performances can be compared against the Forum pre- period as can the publications supported by the Forum. In effect, this leads to a difference-in-difference analysis that controls for some degree of local and global trends when comparing temporal changes between non-Forum and Forum publications by supported investigators. The group of non-Forum publications acts as a counterfactual controlling for differing characteristics of individual awardees. It can be noted that this counterfactual is not perfect, given that the Scopus data on funding acknowledgement has known recall issues. There is a possibility that non-Forum publications are in fact the Forum publications because Scopus coding of funding sources has been imperfect in those cases; or those investigators themselves failed to appropriately mention Belmont support in relevant papers' acknowledgements. Manual spot checks in the non-Forum publications records revealed that such cases are unlikely to be widespread. Conservative interpretations of findings on the Forum publications vs non-Forum publications by awardee investigators would account for all scenarios in this respect.

Secondary comparators using bibliometric and altmetric indicators on peer-reviewed publications: Additional comparators to the CRAs for the peer-reviewed scientific outputs were identified through a data set approach. This approach entailed delineating a global set of publications with similar topics to those of CRA publications (falling mostly in the Science-Metrix categories of Ecology; Environmental Sciences; and Meteorology & Atmospheric Sciences). These publications were retrieved from Scopus using keyword-based queries and citation relationships. Science-Metrix was able to identify other notable funders within this topical data

set using the information from the publications' acknowledgements in Scopus. It should be noted that these comparators were not selected according to their similarity (organisation, mission, or topical focus) to the Forum model, but simply based on their occurrence within the delineated publication set. Finally, the thematic set of comparable publications allows the calculation of world reference performances level that provides a baseline for benchmarking.

Analytical periods: The outcomes and impacts captured through bibliometric and altmetric methods are realised after the start of research projects, sometimes many years afterwards (often after the end of grant support). Therefore, robust bibliometric, altmetric, and network analysis findings were produced with careful consideration of the influence of analytical periods on the observations recorded. Science-Metrix conducted the core bibliometric, altmetric, and network analyses on the outputs of those projects funded by CRAs issued in 2012, 2013, and 2014 (amounting to five CRAs and articles published between 2013 and 2018). For citation-based indicators, only publications from those five CRAs published between 2013 and 2016 were retained, to allow for sufficient outcome realisation periods. For BiodivERsA, publications from the 2008 call and made available between 2009 and 2018 were included (or to 2016 for citation indicators). For all other comparator groups, publications between 2014 and 2018 were used (or to 2016 for citation indicators).

Summary of analyses and roadmap

As is typical in program evaluation, a single bibliometric indicator can contribute to answering multiple evaluation questions (oftentimes, however, deployed in slightly different ways – with different analytical periods, to take just one example). To facilitate navigation of this report, Table below provides an overview of the analyses performed to answer selected evaluation questions, as well as the indicators deployed in each.

Table I Summary of analyses

Evaluation question or narrative function	Table	BF outputs of interest	Dimensions or indicators
Introductory overview	Table 1	Journal publications	Output volume, open access, international co-publication, interdisciplinarity, multidiscplinarity, citation impact (restricted selection)
EQ. 3: Networking effects: "How effective has the Belmont Forum been in generating new collaborations and partnerships across various sectors, disciplines and countries around the globe?	Table 2	Journal publications	International co-publication, authorship, interdisciplinarity, multidisciplinarity, new collaborations
	Table 3	Journal publications	Inter-sectoral co-publication
EQ. 4: Scientific outcomes: To what extent did the Forum contribute to the science base for environmental change (understanding, mitigation and adaptation)?	Table 4	Journal publications	Citation impact
EQ. 5: Wider dissemination of knowledge: How effective has the Belmont Forum been in disseminating knowledge and other outputs generated by the Belmont Forum? To what extent were results of the Belmont Forum disseminated, taken up and discussed beyond academic circles?	Table 5	Journal publications	Open access, altmetrics mentions
	Tables 6 and 7, Figure 1	Non-journal outputs	Output volume, web citations, hyperlinks, usage statistics
EQ. 6: Policy effects and outcomes: To what extent did results of the Belmont Forum foster policy debate or developments at international and national level or facilitate policymaking / implementation?	Table 8	Journal publications	Policy citations
EQ 15: What is the added value of the Belmont Forum (compared to other initiatives at various governance levels)?	Table 9	Journal publications	Recapitulative: international co-publication, inter-sectoral co-publication, interdisciplinary, multidiscplinarity, altmetrics mentions

Source: Prepared by Science-Metrix

Indicator definitions and a detailed methodological is provided in section 0.

Notes for interpretation

Science-Metrix, like other bibliometric academic groups or commercial providers, appraises the meaningfulness of bibliometric findings primarily based on expert judgment, supported by extensive prior experience as well as specific assessments of the individual contexts of programs and their evaluations. Relying on this approach, Science-Metrix analysts make use of rough rules of thumb to pinpoint comparisons between observations of an effect size that reached or exceeded an effect of practical importance.

For example, the minimum threshold for meaningful differences in the ARC, a widely used indicator of scientific impact, is roughly 0.1.[51] In practical terms, if a group's ARC exceeds the

[51] Campbell, D. and Struck, B. (2019). Reliability of Scopus author identifiers (AUIDs) for research evaluation purposes at different scales. *Proceedings of the 17th International Conference on Scientometrics & Informetrics (ISSI)*.

score of another group by exactly 0.1, it means that the former group's papers are 10 percentage points above the latter group relative to the average world paper. Based on our experience gained on many prior evaluation projects, it generally appears difficult for funding programs to induce positive changes larger than the above minimum threshold for established researchers. Such researchers have possibly reached, or nearly reached, a plateau in this regard. Alternatively, it usually appears easier for funding programs to induce positive changes exceeding the above minimum threshold for early-career researchers. Differences between unsupported and supported applicants to funding programs also often reach values larger than the above minimum threshold.

Thus, depending on the context, the normative interpretation of observed differences could be different. In any case, a difference of around 0.1 in ARC will typically be interpreted with care as a slight lead (relative to another group) or improvement (over time). Whenever differences are around the minimum threshold for a given indicator, triangulating data with inputs from the evaluations other lines of evidence is warranted to better appreciate the reliability of the findings.[52] In such cases, the end users of these results (e.g., program stakeholders) should always remain critical of the reported normative interpretation of observed differences, in order to account for contextual information that might have been unknown to those performing the analyses.

In research evaluation—where bibliometric assessments typically rely on random samples of large cohorts of researchers, or on the entirety of a program's cohorts as in this study—the use of formal statistical testing is also warranted to enable statistical inference on future cohorts of the corresponding programs. Where relevant, hypothesis testing was performed to assess the statistical significance of the observed differences and differences-in-differences. Robust *p*-values were estimated using a bootstrapping procedure (unless otherwise stated) instead of traditional statistical testing. This procedure successfully addresses some of the challenges of working with bibliometric data, in a better way than would be feasible with traditional statistical testing (see section 0). Apart from the *p*-values, the bootstrapping procedure was also used to estimate 95% confidence intervals of the bibliometric indicators as well as of differences (or differences-in-differences) between groups. These intervals are sometimes shown in brackets below point estimates in this study's tables. Note that due to space limitations, the differences between groups are not reported in this study's tables as they can be inferred from the scores of the groups being compared. P-values are reported for the difference between the average score for the Belmont set of publication and each comparator. In this study, low p-values suggest that the Belmont papers scores are, on average, higher than the relevant comparator."

The *p*-values reported in this study's tables were highlighted with symbols whenever they are smaller than the customary threshold of significance (*: $p < 0.01$; †: $p < 0.05$; ‡: < 0.1). However, the *p*-values and other statistics should be interpreted while keeping in mind that this report is part of a much larger investigation in which multiple lines of evidence are used. Accordingly, if a *p*-value is larger than 0.05 but smaller than 0.15, we would still argue that the observed difference is likely to be observed in future cohorts of the program, assuming that their characteristics remain roughly unchanged and that the program does not drastically change. In that case, applying the principle of the convergence of partial indicators through triangulation with the evaluations other lines of evidence will help derive robust conclusions.

[52] It is also possible to triangulate findings from different bibliometric indicators relying on a similar information source (e.g., citation data) and yet capturing slightly different phenomena (e.g., scientific impact looking at the overall production [average/global impact] of a group or looking at the most cited papers [excellence]).

It should be noted that findings for comparator groups provided here were computed with a goal to provide comparative reference points for Forum performances only, together representing a variety of funding models in the aggregate against which to appraise the Forum model. Comparison between reference groups are not encouraged and cannot be considered as robust. For instance, while the findings provided could technically allow comparison of ERC and NSF performances, such comparisons would make abstraction of the vastly different scopes and missions of these two agencies.

In the text below, findings from difference comparator groups are sometimes presented in a sequence (for example: "the Forum publications' achievements were measured at value X, and were closely followed by those of funder A and then those of funder B"). Following what was just mentioned, such sequences can only be read as enumerations of discrete comparisons against measurements for the Forum. These sequences can in no way be considered as rankings for multiple funder performances. This is particularly true given that statistical tests on the significance of differences in performances were only performed for comparator versus the Forum observations, and never between comparators.

Finally, the Forum's mission puts a clear emphasis on support for transdisciplinary research. Bibliometric indicators provide insights into multiple, discrete components of transdisciplinary practice, including integration of disciplinary diversity in new research findings (interdisciplinarity); collaborative work that crosses disciplinary boundaries (multidisciplinary); and collaborative work with a strong participatory, local or co-productive orientation. An overall assessment of the Forum achievements in transdisciplinarity can be obtained by combining these discrete components at the time of analysis and interpretation. Additionally, the discrete components of the transdisciplinarity concept dealing with intellectual diversity (interdisciplinarity and multidisciplinary) will sometime be regrouped in the text under the generic term of cross-disciplinarity.

Results

Introductory overview

Key section findings

Generally, the overall (descriptive) picture portrayed from this high-level scan is very positive for the Forum, showing that the funder has certainly made major achievements not only in scientific excellence but also for core missions in supporting transdisciplinarity and North-South collaboration.

- ICR in the Forum articles was 73%, below the observation for the ERC but above the other measurements
- North-South ICR was 44%, slightly above the highest observations for the other funders (at 41%)
- The Forum publications were decisively more interdisciplinary and multidisciplinary than those supported by other funder ($HIP_{10\%}$ of 26%, 9 percentage points above the next highest observation; $HMP_{10\%}$ of 28%, 8 percentage points above the next highest observation).
- The Forum citation impact achievements were comparable to those of the ERC in the thematic data set for some highly cited publications ($HCP_{10\%}$ of 40% for the Forum against 37% for the ERC), but below on other dimensions (ARC of 2.70 for the Forum against 3.19 for

ERC). The Forum was well ahead of other funders besides the ERC for its highly cited publications and roughly on the same level as five other funders for its levels and distribution of citation impact across all publications (CDI dimension).

This section aims to provide a high-level, descriptive scan of the position of publication sets supported by the Forum within a field of global research outputs in key thematic areas of the Forum engagement. This overview provides context; but does not provide the main evidence used in the evaluation of performances of the projects supported by the Forum. The evaluation exercise proper will instead be undertaken in the following sections, using a more restricted but also more rigorously selected set of comparators and reference points. A key difference is that only the results presented in the subsequent sections have been subjected to bootstrapping and statistical testing.

Bibliometric and altmetrics findings in this section, like in the remainder of the report, include only the subset of publications from the Forum CRAs launched early enough for citation windows for most of their publications to have fully elapsed (projects funded by the CRAs for which competitions were held between 2012 and 2014). Science-Metrix also computed the indicators presented in Table II for the overall set of publications supported by the Forum, including those more recent CRAs (data not shown). Findings on most indicators were broadly similar to those reported here, with the exception of the cross-disciplinarity indicators. Levels recorded on those four indicators did see a drop in the Forum lead to other observations, with the greater portion of this change attributable to performances from the Climate2015 CRA.[53]

It would be important to note that output volumes (publication counts) for the Forum CRA 2012–2014 publications (as well as prior and non-Forum publications; and BiodivERsA papers in later sections) are not directly comparable to counts for funder-level categories. Indeed, the recall rate using papers' funding acknowledgement in Scopus is far from 100% and possibly varies across funders. For this reason, the number of papers of comparators in the thematic set could not be compared to that for the Forum. Recall levels between groups were assumed to be different given the divergent methods used to retrieve each data set. Nevertheless, these publications offer convenience samples to assess the performance of comparators on other "relative" indicators that are not (or less) related to size of outputs.

Table II offers an overview through a selection of indicators presented in the remainder of the report. It combines traditional indicators of scientific performance (publication volumes – used as an indication of sample sizes rather than as evaluative evidence – –and citation impact indicators) with indicators that capture some of the dimensions that are at the core of the Forum model : openness (share of publications available in OA – OA%) ; collaboration (international co-publication rate – ICR; multidisciplinarity index – MI; share of highly multidisciplinary publications – $HMP_{10\%}$); and intellectual diversity (interdisciplinarity index – II – and the share of highly interdisciplinary publications – $HIP_{10\%}$).

[53] While climate change science emerged from a historical convergence of diverse disciplinary approaches and tools, currently it is not typically recorded as being highly cross-disciplinary in Science-Metrix' science classification. Publications from the area and their references fall to a great extent within a single subfield, Meteorology & Atmospheric Sciences. Do note that this result is substantiated by other findings in the bibliometric community, where there is currently debate about whether climate change science is still highly cross-disciplinary or whether the historical mix of disciplines that has led to its birth has now been fully institutionalized and can no longer be considered cross-disciplinary (for a summary of related discussions, see Olsen, D. S., Brorstad Borlaug, S., Klitkou, A., Lyall, C., & Yearley, S. (2013). *A Better understanding of Interdisciplinary research in Climate Change*. Oslo.).

Table II Overview of Belmont Forum and other funder performances in the thematic publication set, 2013–2019

Groups	Total N Pubs	OA%	ICR	North-South ICR	II	$HIP_{10\%}$	MI	$HMP_{10\%}$	ARC	$HCP_{10\%}$	CDI
Belmont Forum, CRAs between 2012–2014	371	63.6%	73%	44%	1.24	25.7%	1.92	28.1%	2.70	40.1%	26.4
Comparator groups from the thematic set											
World level	98,812	43.9%	34%	17%	1.09	14.2%	1.25	14.8%	1.33	15.2%	8.4
Selected funders combined	23,658	47.6%	44%	23%	1.09	13.0%	1.33	15.1%	1.81	23.4%	20.1
National Natural Science Foundation of China	6,392	37.4%	35%	33%	1.11	14.4%	1.32	13.1%	1.50	18.7%	14.7
National Science Foundation, US	4,062	55.8%	45%	18%	1.08	13.1%	1.34	16.0%	2.09	28.2%	23.6
European Commission	3,528	54.2%	63%	20%	1.10	14.0%	1.45	18.3%	2.41	31.3%	27.8
European Research Council	487	71.8%	80%	29%	1.14	16.3%	1.45	14.0%	3.19	36.5%	33.3
National Aeronautics and Space Administration, US	1,890	59.9%	49%	22%	1.04	9.7%	1.26	13.2%	2.18	28.7%	22.0
Conselho Nacional de Desenvolvimento Científico e Tecnológico, BR	1,529	41.5%	44%	41%	1.04	10.0%	1.24	14.3%	1.54	18.7%	16.8
National Oceanic and Atmospheric Administration, US	1,464	52.9%	39%	17%	0.91	5.2%	1.07	11.1%	2.01	26.9%	19.9
Natural Environment Research Council, UK	1,431	82.2%	64%	21%	1.09	15.3%	1.50	19.7%	2.28	29.1%	23.9
Chinese Academy of Sciences	1,367	39.6%	43%	40%	1.16	13.6%	1.38	12.8%	1.86	22.0%	18.8
Ministry of Science and Technology of China	1,349	35.6%	37%	35%	1.09	12.7%	1.21	11.4%	1.83	23.2%	18.2
Natural Sciences and Engineering Research Council of Canada	1,299	45.1%	48%	12%	1.07	11.0%	1.19	11.7%	1.77	22.2%	20.2
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, BR	1,040	40.9%	45%	41%	1.03	9.8%	1.29	15.9%	1.74	20.0%	17.1
U.S. Department of Energy	921	61.4%	48%	20%	0.98	11.5%	1.29	16.3%	2.41	30.6%	24.3
Australian Research Council	783	48.8%	58%	21%	1.06	11.6%	1.24	11.3%	2.39	27.8%	27.2
Deutsche Forschungsgemeinschaft, DE	772	51.6%	69%	30%	1.10	13.1%	1.37	15.3%	2.27	31.1%	26.7
Bundesministerium für Bildung und Forschung, DE	755	43.0%	62%	32%	1.08	13.4%	1.37	16.5%	2.37	32.7%	27.9
U.S. Department of Agriculture	753	49.5%	31%	16%	1.10	12.2%	1.29	14.8%	2.03	24.5%	20.5
Fundação para a Ciência e a Tecnologia, PT	604	32.3%	57%	13%	1.08	10.5%	1.40	17.4%	1.55	18.6%	18.3
Agence Nationale de la Recherche, FR	595	52.4%	70%	27%	1.10	12.4%	1.56	19.6%	2.04	25.4%	25.2
Japan Society for the Promotion of Science	557	49.6%	51%	25%	1.04	11.0%	1.22	13.1%	1.16	11.4%	5.9
National Center for Atmospheric Research, US	235	61.0%	40%	17%	0.66	2.0%	0.67	4.3%	1.90	28.6%	20.6

Note: The 20 funders within the thematic publication set were selected based on output volume recorded at an early stage of data preparation. Note that these comparator groups were assembled using convenience sampling based on funding acknowledgement, a method with limited recall. North-South ICR: international co-publication rate with a least one author from a country in OECD ODA list and at least one author from a country not on the list
Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

The full publication set covering the thematic areas of the Forum's CRAs (i.e., thematic keywords drawn from publications from the 5 CRAs with the highest output volume)) was made up of more than 98,000 publications, of which around a quarter (almost 23,000) could be matched to at least one of the 20 funders selected in Table II. Unsurprisingly, some of the world's largest funders were dominating the funders' landscape in this thematic data set: the National Natural Science Foundation of China (NSFC; ~6,000), the US National Science Foundation (NSF; ~ 4,000) and the European Commission (EC; ~3,500, including publications supported by the European Research Council – ERC). China and Brazil were the only two OECD ODA countries included here, with three and two funders respectively.

Making scientific publications available under OA modalities is increasingly perceived as a marker of research engagement towards broad circulation of their findings and improving potential for their societal uptake and public engagement. Groups of publications varied greatly in the extent to which they were published under an OA modality or not, with measurements here ranging from 32.3% to 82%, quite possibly as a consequence of the wide variation in national policies for OA. The aforementioned top score was held by UK Natural Environment Research Council (NERC)-supported papers, with the second rank taken by ERC-supported papers (72%). publications supported by the Forum were made available in 64% of cases, one of the top 5 performances here and above both the world level (44%) and the overall level achieved by the 20 main funders (48%).

ICR measurements ranged from 31% (US Department of Agriculture papers) to 80% (ERC papers). World level here was 34%, while the funders' overall score was 44%. Publications supported by the Forum were to a very high extent written as international co-publications, falling just behind

those of the ERC with a score of 73%. European agencies generally appeared with high rates of international co-publications, including the French Agence National de la Recherche (ANR – 70%) or the German Deutsche Forschungsgemeinschaft (DFG – 69%). In terms of North-South co-publication rate, the Forum stands out (e.g., 44% relative to 23% for the funders) relative to all funders in non-ODA countries and perform similarly to the Chinese and Brazilian funders (noting that being ODA countries, these two countries engage in North-South ICR every time they co-publish with North countries).

Looking at the intellectual integration of diverse disciplinary insights (inferred from the disciplinary diversity of a paper's cited references), the Forum was well above all other funders for its II (1.24) far above world level (1.09) and the main funders combined score (1.09). There was little overall variation on the II with the scores of most funders (16 out of 20) ranging between 1.02 and 1.14.

The Forum also stood out for its share of highly interdisciplinary papers. That is for papers falling in the right tail (the top 10%) of the distribution of paper-level interdisciplinarity scores by subfield. With a score of 25.7% on the $HIP_{10\%}$, this placed the Forum much above the world level (14.2%) and the 20 funders (13.0%). Strikingly, the Forum was in fact the *only* funder considered here with a $HIP_{10\%}$ level substantially above the world level (the ERC and NERC were slightly above at 16.3% and 15.3%). This finding converges with other observations made in the scientific literature that traditional funding mechanisms in fact *actively steer away* from cross-disciplinary research.[54] This finding will be discussed in more details below (see section 0). It can be noted that the difference between observations for the Forum publications overall and CRA 2012–2014 papers is seven percentage points (18.3% for the Forum papers overall, data not shown), a drop that is to be traced back to the issue previously discussed with the Climate 2015 CRA.

The Forum publications also achieved very high levels of cross-disciplinary collaboration looking at the diversity of disciplinary background among a publication's co-authors. The authors of the Forum articles, on average, came from a broader variety of disciplinary research backgrounds than those of papers funded by other agencies (MI of 1.92 vs. 1.33 for main funders; HMP10% of 28.1% vs. 15.1%). The Forum accomplishments of this dimension were above the other notable achievements found in the sample of funders, including vis-à-vis those of papers funded by the ANR ($HMP_{10\%}$ of 19.6%); the Fundação para a Ciência e a Tecnologia (17.4%) or the Brazilian Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (15.9%).The difference between CRA 2012–2014 publications and overall Forum publications was quite large on this indicator, at six percentage points below, due once again to the influence of the Climate 2015 CRA.

A look at three citation impact indicators (average of relative citations – ARC; highly cited publications in the top decile – $HCP_{10\%}$; and citation distribution index – CDI) showed that the research supported by the Forum achieved exceptionally high levels of uptake and visibility within scientific communities. Publications supported by the Forum had the highest share falling within the group of 10% most cited publications in their respective subfields ($HCP_{10\%}$ of 40%). These highly cited papers pulled up the relative average of citations received by the Forum papers to 2.70, below however the top score of 3.19 recorded for ERC-supported papers. The ERC, in fact, appeared as the only funder with performances that surpassed those of the BF on citation impact dimensions. There is a higher concentration of ERC publications, compared to the Forum papers,

[54] Rafols, I. et al. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy*, *41*(7), pp. 1262–1282. doi:10.1016/j.respol.2012.03.015; Bromham, L., Dinnage, R., & Hua, X. (2016). Interdisciplinary research has consistently lower funding success. *Nature*, *534*(7609), pp. 684–687. doi:10.1038/nature18315.

among the broader set of citation deciles above the median (CDI of 33.3 compared to 26.4 for the Forum who came 6th on this dimension). While ERC-funded papers generally achieved a higher citation scores than papers supported by the Forum, a greater fraction of this latter group fell among the most outstanding papers (i.e., the top 10% often used as a proxy for breakthroughs). The ERC appeared to proportionately support fewer highly cited papers ($HCP_{10\%}$ of 37%). Given that the ERC is widely regarded as a funder focusing exclusively on top-flight research, the Forum's performances — and on one dimension superior performances —are to be considered as great achievements for the funder. At any rate, the Forum publications' performances towered above world averages and overall main funders' scores. Performances closer but still a fair bit behind those of the articles supported by the Forum could be found with EC-supported papers (ARC of 2.41 ; $HCP_{10\%}$ of 31.3% ; CDI of 27.8) or BMBF-supported ones (ARC of 2.37 ; $HCP_{10\%}$ of 32.7% ; CDI of 27.9).

To obtain a more robust picture the Forum's performance when supporting research projects, the coming sections will focus on comparisons with carefully matched reference groups and with input from statistical testing.

EQ3: Networking effects: How effective has the Belmont Forum been in generating new collaborations and partnerships across various sectors, disciplines and countries around the globe? To what extent are the funded projects truly co-designed and co-created?

Key section findings

Journal publications supported by the Forum achieved much in terms of international collaboration and cross-disciplinarity (either in terms of knowledge recombination or partnerships). Compared against thematically relevant publications supported by other funders, CRA 2012–2014 publications:

- ranked second for international co-publications with an ICR of 73%, 7 percent points below the highest level recorded (by the ERC);
- achieved a North-South ICR of 44%, 11 percentage points above the next highest score;
- achieved the highest level of knowledge integration between disciplines (II of 1.24 ; 0.10 above the next best performance); it also had the highest share of publications falling within the most interdisciplinary decile in their subfield (25.7%, 9 percentage points above the next best funder);
- reached the highest levels of collaborative multidisciplinarity within their subfield, with a $HMP_{10\%}$ of 28.1%, nearly three times the expected volume. Comparators stood below (BiodivERsA at 20.6%) or much below the Forum on this dimension. Similar findings hold in terms of average multidisciplinarity;
- 22% were written as co-publications between authors from at least three different sectors.

It should be noted that papers published by the Forum awardees prior and concurrently to their Forum-supported papers also score very highly on all these indicators (except for inter-sectoral co-publications for which these groups were not analysed). This indicates that the Forum competitions have been successful in attracting investigators that already engage in international collaboration (globally and for North-South perspective) and cross-disciplinarity. Additionally, the Forum funding clearly enabled these investigators to increase even further their collaborative and interdisciplinary practices.

Output volume data shows that about 9% of the publications of the investigators supported by the Forum (4,116 total) could be linked to the Forum support during the period of support (Table III) – although this number could be affected by the incompleteness of BFGO, Scopus acknowledgement information, or researchers' own acknowledgement practices.

Within the publications supported by the Forum originating with projects financed through the 2012–2014 CRAs, almost 73% of publications were written as international co-publications. ERC-supported publications topped this number with a 80% rate, with BiodivERsA and ANR papers roughly at the same level with scores of 70% (with the Forum's lead to both not statistically significant). The Forum held a significant lead on ICR to NERC (64%), EC (63%) and BMBF (62%) publication sets. The Forum observation was well above world level in the thematic set (34%) and the combined main funders' measurement (44%).

CRA 2012-2014 papers also displayed a much higher ICR score than papers produced by the Forum-investigators prior (50%) and concurrently (i.e., non-Forum papers at 59%) to their Forum awards. The non-Forum counterfactual thus show that the Forum induced a net increase of 14 percentage points (p.p.) for its awardees; the Forum papers increased by 23 p.p. relative to the Forum prior publications, whereas non-Forum papers increased by 9 p.p. relative to this latter group of prior papers. This difference-in-differences (DID) was statistically significant. Average numbers of countries and authors per papers have also been calculated to provide a sort of "red flag analysis" on the intensity of international collaboration within the Forum international co-publications. While maximizing the number of authors and countries per paper is most likely not to an end in of itself for funding support, low findings on these dimensions for publications supported by the Forum could act as signals of caution (obviously taking into account some variance in project configurations, and the fact that not all projects partners systematically contribute to all project publications). On average, CRA 2012–2014 projects involved 10 individual partners and 5 countries (in their proposals). Co-authorship of resulting publications should reflect this diversity, at least to some extent. Additionally, ICR findings do not fully capture the intensity of international collaboration –international co-publications can arise through co-authorships involving only two authors from two countries, a situation, which, if widespread, would not match the Forum's vision for international cooperation.

The Forum publications were comparable to those of many funders on these two dimensions (2.7 countries and 7.3 authors), including BiodivERsA (2.7 countries and 7.1 authors per paper), the EC (2.6 / 7.2) which strongly emphasise international networking through its Framework Programmes (FPs), NERC (2.6 / 8.2), ANR (2.7 / 8.6) and BMBF (2.5 / 8.0). *P*-values calculated for differences between the Forum and these funders were all above 0.05. ERC publications were once again well ahead of others on these dimensions (3.5 / 11.1).

Table III Belmont Forum and comparator achievements on networking effects, 2013–2019

Groups	Total N pubs	ICR	Avg N countries	Avg N authors	North-South ICR	II	$HIP_{10\%}$	MI	$HMP_{10\%}$
Belmont Forum, CRAs between 2012–2014	371	73% [68%|-|77%]	2.7 [2.5|-|2.9]	7.3 [6.5|-|8.1]	44% [39%|-|49%]	1.24 [1.20|-|1.28]	25.7% [21.3%|-|30.0%]	1.92 [1.76|-|2.08]	28.1% [23.8%|-|32.7%]
Matched comparator groups									
Non-BF publications by BF awardees	3,745	59% *	2.5 †	8.2	36% *	1.14 *	15.5% *	1.67 *	21.8% *
BF awardees prior publications	3,044	50% *	2.0 *	6.0 *	27% *	1.15 *	15.6% *	1.71 †	20.7% *
BiodivERsA, 2008 call	426	70%	2.7	7.1	10% *	1.11 *	14.1% *	1.49 *	20.6% *
Comparator groups from the thematic set									
World level	98,812	34% *	1.6 *	4.7 *	17% *	1.09 *	14.2% *	1.25 *	14.8% *
Selected funders combined	23,658	44% *	1.8 *	5.8 *	23% *	1.09 *	13.0% *	1.33 *	15.1% *
National Natural Science Foundation of China	6,392	35% *	1.5 *	5.6 *	33% *	1.11 *	14.4% *	1.32 *	13.1% *
National Science Foundation, US	4,062	45% *	1.9 *	6.1 *	18% *	1.08 *	13.1% *	1.34 *	16.0% *
European Commission	3,528	63% *	2.6	7.2	20% *	1.10 *	14.0% *	1.45 *	18.3% *
European Research Council	487	80%	3.5	11.1	29% *	1.14 *	16.3% *	1.45 *	14.0% *
Natural Environment Research Council, UK	1,431	64% *	2.6	8.2	21% *	1.09 *	15.3% *	1.50 *	19.7% *
Bundesministerium für Bildung und Forschung, DE	755	62% *	2.5	8.0	32% *	1.08 *	13.4% *	1.37 *	16.5% *
Agence Nationale de la Recherche, FR	595	70%	2.7	8.6	27% *	1.10 *	12.4% *	1.56 *	19.6% *

Note: Funders from the thematic set selected to include the 3 largest by output volumes (NSFC; NSF; and EC); and the top 3 by $HIP_{10\%}$ scores. North-South ICR: international co-publication rate with a least one author from a country on the OECD ODA list and at least one author from a country not on the list. One-tail test for differences of means between Belmont Forum and each comparator: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

In terms of North-South collaboration (share of papers containing at least one author from an OECD ODA country and at least one from a non-ODA country), the Forum publications were well ahead of others (44%), with the single funder from an ODA country coming closest (NSFC at 33%). The impact of the Forum mandate was quite apparent here, as neither BiodivERsA or the ERC, which may focus on fostering collaborations within the European Research Area – albeit ERA priority 6 is in indeed about international collaborations beyond it – scored very highly on this dimension (10% and 29% respectively). Additionally, the hypothesis of a specific fostering effect of Forum funding on North-South collaboration is also supported by quite lower share observed for the Forum awardees prior (27%) and concurrently (36%) to the Forum funding (both leads statistically significant).

Considering now the capacity to recombine knowledge from various disciplines within their scientific papers, investigators supported by the Forum again displayed considerable achievements on this dimension. An II of 1.24 was recorded for CRA 2012-2014 publications, well above those of other funders. Closest to this performance were prior articles by the Forum investigators (1.15), non-Forum articles by the Forum investigators (1.14) and BiodivERsA publications (1.11). These findings again show 1) that the Forum calls have been able to select highly interdisciplinary researchers, 2) that the Forum fostered a further increase of their level of interdisciplinarity, and 3) that some of the features of the joint calls (also employed by BiodivERsA model) appear successful in fostering interdisciplinarity relative to less specialized funding models (i.e., relative to national funding agencies with a broader range of funding mechanisms).

The Forum publications also stood out markedly in terms of interdisciplinarity when looking at the tail of highly interdisciplinarity papers (the top 10%) instead of looking at the average. As many as 25.7% of the Forum publications fell within this leading group of highly interdisciplinary papers, more than twice the expected number. The same observations made for the II apply here concerning the Forum's success in selecting highly interdisciplinary applicants ($HIP_{10\%}$ of 15.5% for prior papers by investigators supported by the Forum), in improving their scores (15.6% for non-Forum papers) and the notable performance of BiodivERsA relative to the other comparators of the Forum (14.1%).

The MI dimension captures the collaborative aspect of cross-disciplinary research. It measures the degree of diversity in the disciplinary background of a paper's co-authors. The disciplinary background of an author was assessed by the distribution of its publications across scientific subfields (based on Science-Metrix' journal-based classification). Findings here show once again that the Forum CRA 2012–2014 publications reached the highest intensity on this dimension (1.92). Among the selected funders, ANR followed (1.56), BiodivERsA (1.49), and then EC- and ERC-supported articles (1.45 for both). ERC papers were as multidisciplinary as EC-supported papers taken as a whole (scores of 1.24 for both), while they contained a lower share of highly multidisciplinary publications (14.0% for the ERC compared to 18.3% for the EC). Given that the EC strongly supports networking through the cooperation component of its Framework Programmes (FP), which is different from the ERC component focusing on research excellence, this result is not surprising. Prior and non-Forum publications by investigators supported by the Forum also returned higher scores than those of other funders (1.67 and 1.71). The implications thus remain that the Forum competitions have successfully selected highly multidisciplinary investigators and research collectives, and that the support offered itself contributed to further progression of the Forum awardees on this dimension. Looking at the 10% most multidisciplinary papers, very similar observations were made, with the Forum achievements measured at a share of 28.1% of publications in this case.

Results on inter-sectoral co-publication (ISR) shares are shown in Table IV. As stipulated in the inception report, this indicator was only computed for publications by the Forum awardees given the labour-intensive task of coding author addresses by sector of activity. Looking at basic descriptive statistics on authorship – shares of papers with at least one author in a given sector – the CRA 2012-2014 papers mentioned in a proportion of 96% a university-based author. A share of 43% of publications included at least one author based at a governmental agency – do note that Science-Metrix was not able to accomplish in-depth manual curation of affiliations that would have allowed to distinguish between authors at executive or legislative branches of government (policymakers) and authors from government-funded research center, engaged in what is often called "regulatory science". Based on Science-Metrix analysts' judgement, the vast majority of affiliations in this category appeared to belong to the second group. A share of 37% of the Forum papers saw contribution from authors at large research centers (LRC), a category that includes large networks of government-funded research centers such as the Max-Planck institutes in Germany or the Russian Academy of Sciences institutes. For most intents and purposes these institutions could be grouped together with universities and academia more broadly within this analysis. A share of 29% of the Forum CRA papers had contribution from an author located at 'other research centers (ORC), a category included a dispersed set of independent research groups that could not be clearly classified or characterized. These centers may have been think tanks or research arms of philanthropic organizations, sometimes veering close to non-governmental organizations (NGO), but always with a clear scientific focus. Authors from NGOs appeared on almost 9% of publications supported by the Forum. Authors from inter-governmental organizations (IGO) appeared on slightly more than 5%. Finally, authors from private companies appeared on slightly less than 5% of papers. It should be remembered in interpreting these results that some authors did hold affiliations crossing multiple sectors.

Table IV Belmont Forum achievements in inter-sectoral collaboration, 2013-2018

Authorship	Belmont Forum, CRAs between 2012–2014	Non-BF publications by BF awardees	BF awardees prior publications	BiodivERsA, 2008 call
Descriptive statistics				
Acad	96.2%	94.3%	93.1%	94.8%
Gov	43.1%	48.4%	44.2%	51.9%
LRC	37.1%	37.4%	32.5%	52.8%
ORC	29.0%	23.8%	17.9%	29.3%
NGO	8.7%	4.3%	2.8%	3.5%
IGO	5.4%	5.5%	3.3%	1.2%
Priv	4.6%	8.6%	5.9%	7.0%
Collaborations within research-oriented sectors				
Acad & LRC	34.4%	33.8%	29.1% †	50.2%
(Acad.LRC) & Gov	41.5%	46.5%	40.7%	49.5%
(Acad.LRC) & ORC	27.1%	22.9% †	16.5% *	28.4%
Broad collaborations				
(Acad.LRC.ORC.Gov) & (NGO.IGO.Priv)	15.4%	15.1%	10.6% *	11.0% †
(Acad.LRC.Gov.ORC) & NGO	8.7%	4.3% *	2.7% *	3.5% *
(Acad.LRC.Gov.ORC) & IGO	5.4%	5.4%	3.2% †	1.2% *
(Acad.LRC.Gov.ORC) & Priv	4.3%	8.6%	5.9%	7.0%
(Acad.LRC) & (ORC.Gov) & (IGO.NGO.Priv)	3.5%	4.5%	2.2%	1.2% †
(Acad.LRC.ORC.Gov) & NGO & IGO & Priv	0.3%	0.3%	0.1%	0.0% ‡

Note: Shares of publications with a least one author with an affiliation in the category or categories of interest. Do note that authors may have multiple affiliations falling within multiple sectors. Acad: universities and academic institutions. Gov: government research centers (mostly) and governmental agencies. LRC: Large research centers, often government-funded and quasi-academic, such as the Max-Planck network of institutes. ORC: Other research centers, whose institutional status could not be clearly established but which appeared to be independent research centers; charity-based; or (less often) government-funded. IGO: inter-governmental organisation. NGO: non-governmental organisation. Priv: private. Sectors regrouped within parentheses were interchangeable in the query; '&' signs denote a necessary combination. One-tail test for differences of means between Belmont Forum and each comparator: * p<0.01, † p<0.05, ‡ p<0.1. Significance testing was not performed on descriptive statistics.
Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

In terms of the shares of publications supported the Forum that can be counted as inter-sectoral co-publications, university affiliations were taken as starting points given their aforementioned prevalence, and given the expectation that non-academic partners, even when they publish peer-reviewed publications, seldom do so without any contribution from an academic partner. Publications were then examined to determine whether they contained a co-publication between a university-based author and at least one author from another sector. As defined above, the LRC, governmental and ORC categories were all composed of affiliations with a strong research orientation, despite being outside of the higher education environment narrowly defined. Given this, is it not surprising to find the highest ISR figures in this subset of collaborations. the Forum publications contained university or LRC collaborations with governmental research centers or agencies in almost 42% of cases, and with ORC-based authors in 27% of cases.

Interestingly, the Forum support does appear to have fostered co-authorship with ORC partners, with a statistically significant increase over prior practices of the Forum investigators (measured at almost 17% of papers). This increase was also above that seen in parallel publications by the Forum investigators. The score recorded for the Forum publications was below that of BiodivERsA publications, however. the Forum publications were also not significantly more likely to have been written in collaboration with regulatory scientists or governmental partners than in the Forum investigators' prior papers, whereas the same figure had gone up in their concurrent

publications. Again, BiodivERsA also displayed a higher share of publications written as a result of this type of collaboration (50%).

The Forum publications were also examined to identify publications that combined multiple sectors, and especially the three non-research sectors included in the analysis. Taking a loose filter that accepts collaborations from any of the research sectors together with any of the non-research sectors ("(Acad.LRC.ORC.Gov) & (NGO.IGO.Priv)" in the table), the ISR share was slightly more than 15% of the Forum publications. This observation was higher than for prior publications by the Forum investigators (11%, with a statistically significant difference) and biodiverse (also 11% with a statistically significant difference), but roughly the same level as found in concurrent publications by the Forum investigators (15%).

Looking at more narrow filters of collaboration, it will become clear that the Forum publications collaboration shares are driven by NGO participants, whereas those of concurrent publications by the Forum investigators are instead driven by partnerships with private sector authors. The Forum collaborative publications appear to have included NGO-based authors to high levels (almost 9%; operationalized as "(Acad.LRC.Gov.ORC) & NGO"), significantly above parallel publications by the Forum investigators (4.3%), prior publications (2.7%) and BiodivERsA (3.5%). Higher shares of the Forum publications were also written in collaboration with an IGO-based author (5.4%) than publications prior to funding (3.2%) or BiodivERsA publications (1.2%). Concurrent publications by the Forum investigators saw the same share of these publications with an IGO-based co-author, however. The Forum funding may not have been the deciding factor in fostering this specific type of co-publication for supported investigators, although the the Forum competitions were successful in identifying investigators with good potential in this respect. Looking at co-publication with authors based in the private sector, the Forum publications recorded a lower level on this dimension than prior publications by the Forum investigators (4.3% to 5.9%). By contrast, the share of such co-publications increased in parallel papers by the Forum investigators compared to the figure in prior articles (8.6%). BiodivERsA articles also showed a higher share of these co-publications (7.0%). It appears that the Forum support has contributed towards a shift in the focus of collaborative activity for supported investigators, who may have moved away from private sector partners to NGO-based collaborators. This shift, however, was not accompanied with a net increase in collaborative activity overall, since, has already shown above, the aggregate figure for collaborative activity irrespective of the exact sector (NGO, IGO or private) was roughly the same for the Forum publications and concurrent papers by the Forum investigators.

Turning to more intensive collaborations, publications that included affiliations from at least three different sectors, one of which would be non-research oriented ("Acad.LRC) & (ORC.Gov) & (IGO.NGO.Priv)") were quite sparse in all comparator groups. The Forum publications were found to hold a share of 3.5% such papers, below the share found in concurrent articles (4.5%). Both sets had share above that found in prior publications (2.2%; non-significant lead for the Forum publications). BiodivERsA publications' share was the lowest at 1.2% (with a significant lead of the Forum publications against this group). Using a very discriminating filter that required the presence of at least one co-author from all "non-research sectors" yielded a very low number of observations, showing that the Forum projects did not achieve the most intensive forms of collaboration in the writing of peer-reviewed publications, but that neither did projects from comparator groups.

The evaluation question that framed the analyses presented in this section was: How effective has the Forum been in generating new collaborations and partnerships across various sectors, disciplines and countries around the globe? To what extent are the funded projects truly co-designed and co-created?

The overall picture that emerges in answer to EQ3 was that, and considering only peer-review publications as project outputs, yes, the Forum has been efficient in generating new collaborations and partnerships across actors, disciplines and countries. The Forum articles were by far the most inter- and multi-disciplinary in the comparative analysis conducted here. They also displayed the highest proportion written as North-South co-publications. They recorded the highest proportion to include authors based with an NGO, and high proportions of publications written with an IGO-based author.

The overall picture provided by the observations reviewed here concerning the secondary question is that, again when considering only peer-reviewed publications, academic partners located in institutions in the "North" remain the core group to contribute to the production of these research outputs. An important limitation of the analysis here is that it can in principle be expected that co-designed and co-designed work in transdisciplinary research may not in fact lead to peer-reviewed publications as their main outcomes. It may also be possible that a broad swath of stakeholders would have co-designed a research project without participating in the writing of peer-reviewed publications. While section 0 below reviews non-journal outputs of the Forum projects, this analysis could not evaluate the degree of disciplinary, sectorial and country diversity in the same way it was done here for peer-reviewed publications.

EQ4: To what extent did the Forum contribute to the science base for environmental change (understanding, mitigation and adaptation)?

Key section findings

For EQ4 "to what extent did the Forum contribute to the science base for environmental change (understanding, mitigation and adaptation)?", findings show publications supported by the Forum to have had high levels of citation impact within relevant subfields. The citation impact achievements of the publications supported by the Forum were often at a level well above those of comparators, and sometimes close or above those of ERC-supported articles – a funder widely recognized for its focus on scientific excellence.

- A share of 40% of publications supported by the Forum were within the world's 10% most cited publications in their corresponding subfield (the ERC recorded a share of 36.5%, with the the Forum lead not being statistically significant);
- Publications supported by the Forum were tendentially published in prestigious journal, with a WCS observed at 1.74 (ERC scored 1.76); and
- Altogether, the Forum papers achieved high impact with a CDI of 26.4. A score much above world level and similar to other funders in the set of selected comparators (i.e., BiodivERsA, BMBF and EC), although ERC publications were well ahead those of others on this dimension (CDI of 33.3).
- The Forum successfully selected highly influential scholars and likely enabled further improvements to their scientific impact.

Citation-based indicators have long been used as proxies of the impact (or influence) of scientific publications in assessing the value of their contributions to an expanding knowledge base. Nevertheless, it is important to consider that citation-based indicators as a whole rely on the assumption that citations are generally used to express intellectual debt, to point to the prior work on which one is building in generating research questions, observations, or methods to take a few examples. However, citations are used for other purposes as well; in fact, citations are sometimes used to formulate critiques, which conveys the opposite of the positive ascription that

is tacitly assumed here to citations. Citation-based indicators rely on the notion that critique-driven citations are much less frequent than impactful citations, and therefore of negligible influence on bibliometric study results, so long as one is working with sufficiently large numbers.

Of the 371 CRA 2012-2014 publications, 157 were published early enough for their citation window (i.e., the period over which they accumulated citations) to be long enough to allow for robust citation metrics. Citation-based indicators presented here (with the exception of the Weighted CiteScore – WCS) were based on this subset of 157 publications (Table V).

Table V Scientific impact of Belmont Forum contributions in expanding the science base for environmental change, 2007-2019

Groups	Total N pubs	N pubs citable	ARC	$HCP_{10\%}$	$HCP_{1\%}$	CDI	CDC	WCS
Belmont Forum, CRAs between 2012–2014	371	157	2.70 [2.2|-|3.3]	40.1% [32.7%|-|47.7%]	6.4% [2.5%|-|10.2%]	26.4 [22.2|-|30.6]		1.74 [1.60|-|1.90]
Matched comparator groups								
Non-BF publications by BF awardees	3,745	2,195	2.09 †	24.5% *	4.4%	19.6 †		1.45 *
BF awardees prior publications	3,044	2,755	2.26 ‡	27.3% *	4.2%	19.8 †		1.43 *
BiodivERsA, 2008 call	426	405	2.26 ‡	29.9% †	4.2%	27.5		1.61 ‡
Comparator groups from the thematic set								
World level	98,812	48,782	1.33 *	15.2% *	1.7% †	8.4 *		1.15 *
Selected funders combined	23,658	10,038	1.81 *	23.4% *	3.0% †	20.1 †		1.39 *
National Natural Science Foundation of China	6,392	2,511	1.50 *	18.7% *	1.8% †	14.7 *		1.20 *
National Science Foundation, US	4,062	1,835	2.09 †	28.2% *	4.7%	23.6		1.56 †
European Commission	3,528	1,463	2.41	31.3% †	5.4%	27.8		1.54 *
European Research Council	487	185	3.19	36.5%	10.7%	33.3		1.76
Natural Environment Research Council, UK	1,431	774	2.28 ‡	29.1% *	5.0%	23.9		1.64
Bundesministerium für Bildung und Forschung, DE	755	296	2.37	32.7% ‡	3.8%	27.9		1.48 *
Agence Nationale de la Recherche, FR	595	262	2.04 †	25.4% *	3.1% ‡	25.2		1.56 †

Note: N pubs citable: Number of articles in the comparator group's publication set for which the minimal citation window for computing robust citation indicators has been reached and for which citation-based indicators can be computed. One-tail test for differences of means between Belmont Forum and each comparator: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus database (Elsevier)

Overall, the Forum CRA 2012–2014 publications reached great levels of citation impact and were consistently above those of most other funders considered. In terms of average citation impact, the Forum publications recorded an ARC of 2.70, only behind the ERC (3.19) and above the BMBF (2.37) and EC (2.41) publications (with non-significant leads). The Forum measurement on the ARC dimension was significantly above the remaining observations included in Table V.

The CDI, another measure of impact accounting for all publications which is less sensitive than the ARC to highly cited publications, indicated that the Forum publications (26.4) are comparable to a broader set of funders. The ERC again had the highest score here (33.3) followed by the BMBF (27.9), EC projects (27.8), and BiodivERsA (27.5). The difference in the placement of the Forum based on the ARC and the CDI is attributable to the strong scores of the Forum's highly cited publications which pull the ARC up. Given the small number of CRA 2012–2014 papers that could be used in computing these metrics, it is our view that the CDI provides a better reflection of "average" performance here.

The high CDI of the Forum — and of ERC, BiodivERsA, BMBF and EC — shows that they overall have a very strong research influence. For example, there is a strong concentration of CRA 2012–2014 papers in the highest citation deciles with a majority of them (64%) falling in the top three deciles in which the share is always above expectations (as revealed by the green bars to the right of the CDC [a visual depiction of the CDI]), especially in the top 10%. This comes with a

much smaller share of papers than expected in the three least citation deciles (6% vs. 30% expected; revealed by the long red bars to the left of the CDC).

Focusing on very highly cited publications (i.e., $HCP_{1\%}$) as a reflection of contribution to scientific excellence/breakthroughs, findings similar to the ARC are obtained. A share of 6.4% of the Forum publications fell within the exceptional group of 1% most cited publications, putting the funder most likely only behind the ERC with a proportion of 10.7%. the Forum's score was non-significantly above those of the EC (5.4%), the NERC (5.0%), the NSF (4.7%), BiodivERsA (4.2%), and BMBF (3.8%) publication sets. Based on $HCP_{10\%}$, the Forum publications came out first by contrast, with a proportion of 40.1%, a larger share than in the ERC publication set (36.5%, the *p*-value to reject the null hypothesis that the Forum is smaller or equal to ERC was equal to 0.237). These diverging leads between the Forum and ERC on the $HCP_{10\%}$ and $HCP_{1\%}$ dimensions show that within the leading group of highly cited papers (top 10%), the distribution remains skewed in favor of ERC (for its publications among the top 1% most cited) which helps explain its much higher ARC.

The Forum publications also tended to be published, on average, within the most prestigious journals in the comparison group (WCS of 1.74), together with the ERC (1.76). the Forum's lead to BiodivERsA on this dimension was meaningful (0.13 points) but slightly uncertain, being associated with a *p*-value of 0.075.

It should also be noted that papers produced by the Forum-investigators prior and concurrently (i.e., non-Forum papers) to their the Forum awards often reached significantly lower citation impact levels than the Forum publications (with exceptions for the $HCP_{1\%}$; and the ARC for prior publications). While prior and non-Forum publications by the Forum investigators displayed measurements below those of the majority of the selected funders in Table V, their scores remain well above world level. This points to the Forum being successful in selecting outstanding scientists through it CRA calls in addition to making a real contribution towards building capacity for high citation impact research; for example, statistically significant, and positive, difference-in-differences (i.e., difference between observed change from prior to the Forum papers relative to observed change from prior to non-Forum papers) were observed for $HCP_{10\%}$ (+15.6 p.p.), CDI (+6.8), and WCS (+0.30).

When considering findings from all citation impact indicators taken together, it is clear that the Forum-funded projects have made highly impactful contributions to the scientific communities working on environmental change and associated topics. Overall, it has been second only to the ERC in the panel of comparators considered here, and placed well ahead of the other groups, a position that is enviable. Given the possibility that a focus towards transdisciplinarity, societal engagement and co-production may in principle come at a trade-off with scientific excellent, the effectiveness and impact found on this dimension can be concluded to have reached a level that is above and beyond expectations.

EQ5: Wider dissemination of knowledge: How effective has the Belmont Forum been in disseminating knowledge and other outputs generated by the Belmont Forum? To what extent were results of the Belmont Forum disseminated, taken up and discussed beyond academic circles?

Bibliometrics and altmetrics conducted with journal-based publications provide evidence on the broader resonance of these classical scientific outputs solely. However, a heavy emphasis with the the Forum programme was also put on the production of local, co-productive and participatory outcomes. Tentative quantitative evidence on these non-journal outputs are provided through custom altmetrics indicators. As altmetric methods are far more developed for traditional publication outputs (Section 0), their customised application to the non-journal

outputs promoted by the Forum are treated in a different sub-section (Section 0). Note that uptake in policy, because it is sourced from a distinct data source, is treated in Section 0 even if it is part of altmetric methods.

Online attention towards journal-based outputs

Key section findings

Data on the share of OA publications and on the extent of altmetric mentions (in journalistic news, Wikipedia, Twitter and Facebook) provided clues as to the degree to which scientific publications supported by the Forum have been taken up and discussed by a broad public and range of potential users extending beyond academic circles. The Forum publications were found to have:

- The highest observations among comparators for news and Facebook mentions (with 24.2% and 21.3% scores, respectively, on the $HAP_{10\%}$ indicator). This suggest that the CRA funding model leads to comparatively high societal impacts.
- High, but not the highest, measurements on OA% (63.6%), Wikipedia and Twitter mentions (12.4% and 25.1% $HAP_{10\%}$ scores, respectively).
- Using a counterfactual, the Forum funding also appears to have promoted an increased uptake of the outputs of its awardees in the news, Twitter and Facebook, thereby likely contributing to an increased societal impact of their research outputs. A similar effect was very likely for Wikipedia mentions as well, although statistical robustness of findings was not definitive here.

The first indicator examined to answer these questions is the share of the Forum publications available under an Open Access modality (Table VI). This indicator is not a direct measurement of outcomes in terms of user engagement beyond academic circles. It is, however, a process indicator that captures the realisation of one of the privileged mechanisms currently in use to foster open science, wider circulation of findings, and sometimes even research user engagement.

CRA 2012-2014 publications were published under an OA modality in a proportion of 63.6%. This observation was below levels recorded for publications supported by the NERC (82.2%), BiodivERsA (73.6%), and the ERC (71.8%). The Forum's score was above or well above the remaining funder publication sets, and also well above world level and the main funders' combined figure. The Forum support appeared to have moderately increased awardees' propensity to publish with an OA license (scores for papers published by the Forum awardees prior and concurrently to their papers supported by the Forum, being seven and six percentage points below the the Forum figure).

Table VI Belmont Forum achievements in disseminating knowledge and attracting online attention, 2007–2018

Groups	Total N pubs	OA%	News			Wikipedia			Twitter			Facebook		
			AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$	AMI	$HAP_{10\%}$	$HAP_{1\%}$
Belmont Forum, CRAs between 2012–2014	371	63.6% [59.0%|-|68.9%]	3.60 [2.89|-|4.31]	24.2% [20.3%|-|28.1%]	6.8% [4.3%|-|9.4%]	2.56 [1.39|-|3.90]	12.4% [10.7%|-|14.4%]	2.1% [1.1%|-|3.3%]	1.61 [1.48|-|1.73]	25.1% [21.0%|-|29.5%]	4.9% [2.7%|-|7.3%]	1.98 [1.63|-|2.33]	21.3% [17.4%|-|25.4%]	4.0% [2.2%|-|6.2%]
Matched comparator groups														
Non-BF publications by BF awardees	3,745	57.1% *	1.87 *	13.9% *	2.4% *	1.63 ‡	11.0% ‡	1.6%	1.39 *	19.6% *	3.0% ‡	1.50 *	14.8% *	1.8% †
BF awardees prior publications	3,044	58.1% †	1.43 *	10.7% *	1.6% *	1.25 †	10.8% ‡	1.2% ‡	1.25 *	12.0% *	1.2% *	1.29 *	11.0% *	1.2% *
BiodivERsA, 2008 call	426	73.6%	2.09 *	12.6% *	2.8% *	3.30	15.7%	4.0%	2.63	27.3%	3.2%	1.74	15.1% *	1.4% †
Comparator groups from the thematic set														
World level	98,812	43.9% *	0.91 *	9.5% *	0.9% *	0.95 †	9.9% *	0.9% †	0.93 *	9.7% *	0.9% *	1.01 *	10.1% *	1.0% *
Selected funders combined	23,658	47.6% *	1.41 *	12.1% *	1.6% *	1.29 †	10.5% †	1.3% ‡	1.23 *	14.3% *	1.4% *	1.27 *	12.5% *	1.3% *
National Natural Science Foundation of China	6,392	37.4% *	0.41 *	6.9% *	0.3% *	0.28 *	8.9% *	0.3% *	0.62 *	2.3% *	0.3% *	0.50 *	4.1% *	0.4% *
National Science Foundation, US	4,062	55.8% *	2.82 †	19.5% †	3.4% †	2.29	12.0%	2.5%	1.68	24.2%	2.9% ‡	1.59 †	16.1% *	1.7% †
European Commission	3,528	54.2% *	1.54 *	12.8% *	2.0% *	1.54 ‡	10.9% ‡	1.6%	1.44 *	19.0% *	2.2% †	1.77	17.9% ‡	1.2% *
European Research Council	487	71.8%	2.94 ‡	20.4% ‡	5.2%	3.21	13.2%	3.3%	1.89	32.7%	5.2%	2.08	21.0%	3.0%
Natural Environment Research Council, UK	1,431	82.2%	3.26	20.8% ‡	4.1% †	2.76	12.8%	2.7%	2.08	40.6%	4.5%	1.95	19.2%	1.5% †
Bundesministerium für Bildung und Forschung, DE	755	43.0% *	1.57 *	12.9% *	2.5% *	1.33 ‡	10.3% †	1.4%	1.25 *	14.4% *	1.6% *	1.61 †	15.1% *	1.1% *
Agence Nationale de la Recherche, FR	595	52.4% *	1.11 *	10.6% *	1.5% *	1.76	11.4%	2.0%	1.52	19.4% †	2.0% †	1.56 †	15.3% *	1.7% †

Note: OA%: Share of papers made publicly available under an open access publication mechanism. One-tail test for differences of means between Belmont Forum and each comparator: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Scopus and PlumX databases (Elsevier)

Moving on to altmetric indicators, a panel of four dimensions were retained from the PlumX database: mentions in journalistic news items (with a known bias towards English-language and Northern sources); references to the publications made on Wikipedia; Twitter mentions; and posts on Facebook. Because altmetric mentions often remain rare events, these forms of uptake are generally quantified in terms of share of publications with at least one mention, rather than by rating publications based on the volume of each type of mentions they received as with the ARC, CDI and HCP indicators using scientific citations.[55] Of course, these shares must be normalised by year to account for the varying periods over which papers published in different years have accumulated citations. This process leads to the altmetric mentions index (AMI) whose reference level (i.e., the world in a broader thematic data set) equals 1. Like for the ARC, scores above one denote performance above world level whereas scores below one mean the opposite.

While further details are provided in the annex, it must be kept in mind that altmetric measurements suffer from known limitations. Altmetrics capture online attention or even buzz towards scientific articles, but it is not always clear to which extent public engagement of the sort is meaningful and provokes sustained reflection. This is particularly the case for Twitter mentions.[56] It would be excessive to expect altmetrics to capture deep knowledge transfer or social change processes resulting from research.[57] Additionally, at least a portion of social media and wiki activity is also fuelled by uptake by other scientists, rather than the general public. Research teams themselves often use social media to promote their own journal publications. In the case of the last limitation, however, the use of high activity metrics complementary to the AMI helps to parse away publications whose altmetric score would only reflect self-promotion (e.g., papers with a single mention in Twitter coming from the paper's authors). Here, an approach equivalent to the computation of the HCP indicators for scientific impact is used to produce such metrics, namely the share of highly altmetric-mentioned publications (HAP). Two versions of the HAP indicator are produced, one for the share of papers falling in the 10% of papers with the largest number of altmetric mentions ($HAP_{10\%}$) and one for those papers falling in the top 1% ($HAP_{1\%}$).

Looking first at mentions in journalistic news items, the Forum publications achieve the highest levels of attention within the panel of comparators. The AMI, the relative share of publications mentioned at least once in news items, was measured at more than three and half times (3.60) the share in the overall thematic publication set, while NERC and the ERC followed at 3.26 and 2.94 and (both the Forum leads being non-significant). The share of the Forum publications that have been highly mentioned in news items ($HAP_{10\%}$) is 24.2%, more than twice the expected figure. Again, ERC and NERC accomplishments on this dimension were close behind (20.4% and 20.8%, with the Forum leads significant). The comparison yielded roughly the same patterns for exceptionally highly mentioned publications ($HAP_{1\%}$), with the Forum share of 6.8% against 5.2% and 4.1% for the ERC and NERC (again significant lead in the comparison to NERC only). On all three indicators, the Forum shares were well above the world level and the main funders' combined score. It can also be noted that prior and non-Forum publications by the Forum investigators recorded news mentions levels that were well below those for CRA 2012-2014 publications, leading to statistically significant differences-in-differences for these indicators.

---

[55] Thelwall, M. (2016). *Web Indicators for Research Evaluation: A Practical Guide*. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 8. Morgan & Claypool Publishers LLC. doi:10.2200/s00733ed1v01y201609icr052.

[56] Tahamtan, & Lutz Bornmann, Altmetrics and Societal Impact Measurements: Match or Mismatch? A Literature Review.

[57] Pulido, C. M., Redondo-Sama, G., Sordé-Martí, T., & Flecha, R. (2018). Social impact in social media: A new method to evaluate the social impact of research. *PLOS ONE*, *13*(8), p. e0203117. doi:10.1371/journal.pone.0203117.

Accordingly, the results suggest a potentially strong, and positive, effect of the CRA funding mechanisms on the societal impact (inferred from mentions in news items) of the Forum outputs.

Turning to Wikipedia citations, a proxy for educational impact, achievements of the publications supported by the Forum on this dimension appeared more modest (12.4% on the $HAP_{10\%}$), although it could be noted that the range of observations recorded was quite narrow: all scores were contained within a range extending from 8.9% for the NSFC to 13.2% for the ERC, all close to the world level at 10.0%. BiodivERsA made the one distinctive performance here (15.7% on the $HAP_{10\%}$). Patterns in comparative achievements were very much similar across the three indicators on this dimension. It can be hypothesized that Wikipedia mentions are comparatively rare events and that their distribution was much sparser that it was for news mentions. Consequently, the results presented should perhaps be interpreted in a more favorable light than what would be expected given modest effect sizes. Finally, the specific contribution of the Forum funding to awardee accomplishments amounted to a difference of roughly one and a half percentage points for the $HAP_{10\%}$ in a comparison to prior and parallel publications produced by the Forum awardees, indicating a very likely (*p*-value below 0.1) specific fostering effect for the Forum support. This effect is clearer on the AMI, where the difference-in-difference can be measured at 0.93 (*p*-value again below 0.1).

The Forum publications' recorded mixed findings, depending on the chosen indicator, for their outcomes on the plane of Twitter "buzz", ranging from moderate to very high in terms of placement relative to selected funders. The bulk of the Forum publications appeared to receive somewhat average levels of attention relative to the selected group of funders while this finding co-existed with a disproportionally large set of the Forum publications registering the highest levels of attention. The share of publications receiving at least one mention was 60% above world level for the Forum (AMI = 1.61), a level below that of BiodivERsA (2.63), NERC (2.08), ERC (1.89) and NSF (1.68) publications. On the $HAP_{10\%}$, the Forum articles (25.1%) were also below many other comparator groups: NERC (40.6%); ERC (32.7%); and BiodivERsA (27.3%). However, on the $HAP_{1\%}$, which again captures the share of publications with exceptional levels of attention on Twitter, the Forum publications registered the second highest observation (4.9%), just below the ERC (5.2%). They were followed closely by NERC (4.5%) articles (the Forum lead being non-significant). The strong score for the Forum on the $HAP_{1\%}$ suggests that where the Forum papers stood out on this dimension cannot be purely attributed to self-promotion as can often be the case with Tweets. For all three indicators, the Forum funding appeared to lead to specific increases in awardee realisations on this dimension. For example, an increase of 13 p.p. on $HAP_{10\%}$ (relative to prior publications by the Forum grantees) and a difference-in-differences of 6 p.p. on using concurrent publications by the Forum awardees as a counterfactual were observed (with a *p*-value below 0.01).

The Forum publications found high levels of attention on the Facebook platform. The publication set recorded the highest observations on the three indicators computed. On the AMI, its normalized share of publications mentioned on Facebook was twice the baseline for the overall thematic set (1.98), an achievement similar to observations from the ERC (2.08), the NERC (1.95), and BiodivERsA (1.74 – the Forum leads were non-significant for these three comparisons). The Forum publications' led on the $HAP_{10\%}$ (21.3% with the NERC following at 21.0% – a difference that was not statistically different) and the $HAP_{1\%}$ (4.0% with the ERC following at 3.0%, lead not statistically different). The Forum funding here again may have contributed to raising awardees' performances. To take just one example, the Forum papers had a significant lead of six and ten p.p. relative to parallel and prior papers by the Forum awardees on the $HAP_{10\%}$ indicator.

Online attention towards non-journal outputs

Given the general challenges currently facing quantitative assessments of societal outcomes of research, as well their intermediaries, Science-Metrix has opted to rely on a mixed methods approach combining multiple sources to obtain a better portrait of the Forum achievements on these dimensions.

Strategies and data sources to be triangulated included:

- BFGO self-reported data on project outputs, including non-journal outputs
- Project websites retrieved by a manual online scan
- Additional project outputs retrieved from project websites
- Data on hyperlinking activity towards online non-journal outputs (provided through a commercial web analytics solution)
- BFGO self-reported data on project outcomes
- BFGO self-reported data on stakeholders engaged in the projects

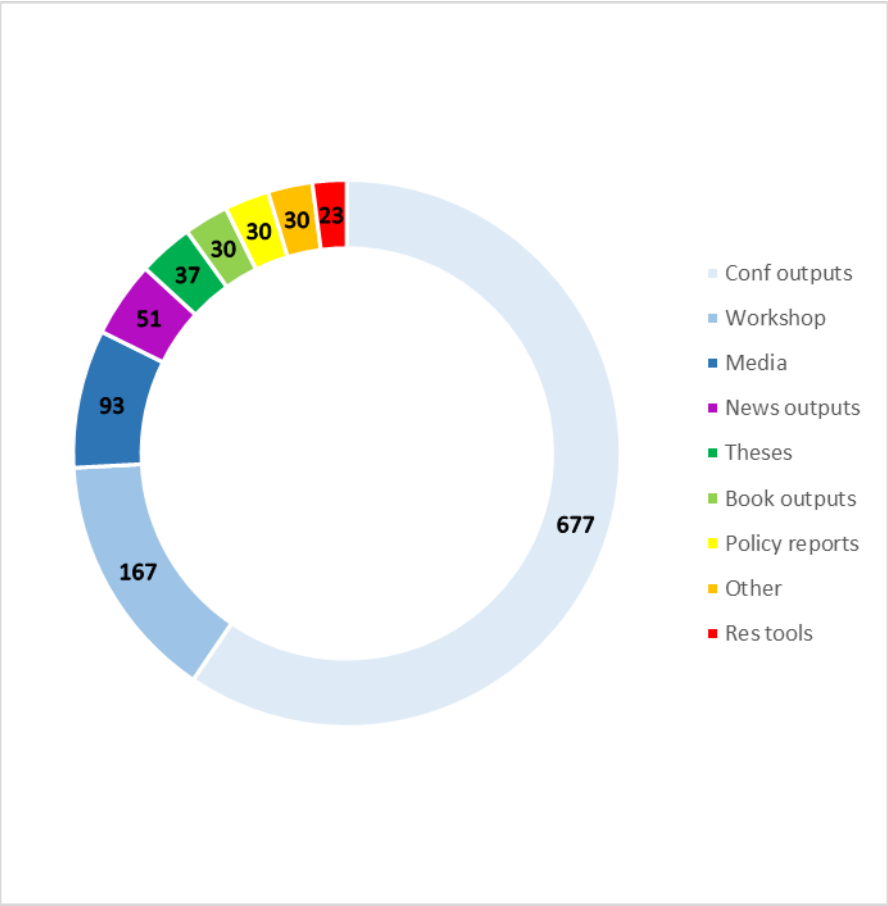
Reviewing these sources and the findings obtained, prospects for triangulation were greatly diminished when it became clear that the level of attention afforded to project reporting varied greatly from one project to the other; and that project managers tended to focus on some of the output and outcomes types mentioned above at the expense of others. For instance, some projects had provided a great amount of details on the journalistic news items written on their research, but nothing on stakeholders and other outcomes. Emphasis was on stakeholders, instead, in reporting from other projects.

Descriptive statistics

Restricting the analysis to project outputs that could not be matched to the Scopus database, a total of 1,138 additional output entries were retrieved. These outputs were produced by projects funded through the CRAs 2012 to 2016.[58] They fell into nine broad categories, presented in Figure . The most voluminous (677) of these classes grouped together various conference presentations, including: talks, poster presentations, and other verbal dissemination activities of an academic character. The next largest category (167 records) was made up of various workshops and discussion activities in engagement, inter-sectoral or practical settings. These could include sessions to disseminate findings to policymakers and practitioners. The third category (titled 'media' in Figure , and counting 93 entries) conflated social media posts, video productions (posted on YouTube or Vimeo) and blog posts. The journalistic news outputs category included 51 items. It should be noted that it included both pieces written directly by the Forum-funded investigators (i.e., op eds); but also scientific journalism making use of findings from a project or of interview material with an investigator.

[58] Although note that BFGO does not contain record for CRA 2012 – entries here were retrieved from project websites only.

Figure 1 Belmont Forum achievements in fostering non-journal and non-academic outputs, 2012–2020



Note: Conf outputs include conference presentations, poster presentations, session organization and other conference-related activities. 'Media' includes social media and video (YouTube, Vimeo) outputs. The 'other' category included working papers, confidential reports, newsletter issues. 'Res tools' includes publicly available data sets, modelling tools and other research tools.

Source: Prepared by Science-Metrix using the BFGO database and web queries

The remaining five categories all included between 37 and 23 records each. Outputs in three of these categories were more targeted towards academic audiences, such as master's and doctoral theses (37 items); book outputs (including monographs, chapters and editing of collective volumes; 30 entries) and research tools (including data sets made available online or modelling tools and components; 23 entries). The 2012-2016 CRA project teams had collectively published or contributed to 30 policy reports. The 'other category' included working papers, confidential reports, newsletter issues and entries on new research projects spun-off from projects supported by the Forum.

In addition to these academic and non-academic outputs, BFGO also record self-reported entries on project stakeholders and outcomes. For CRAs between 2013 and 2016, 57 projects reported 201 outcomes; 63 projects reported 514 stakeholders. These numbers should not be used in an evaluative fashion, as markers of productivity, because definitions of a single unit of stakeholders or outcomes appear to have varied greatly between projects. Of particular interest to this evaluation, project reports included a categorization of these outcomes as 'economic effect'; 'policy effect'; 'social effect'; 'societal effect'. These four categories taken together accounted for 73 records.

Web citation impact

It can be noted that findings from the web citation analysis could not be normalized by year and discipline. The temporal factor, in particular, could have played a role in producing differential web citation impact levels between the Forum and BiodivERsA outputs, given how BiodivERsA outputs were most likely to have been produced between 2008 and 2015, while the Forum outputs were produced later. Additionally, online dissemination practices have considerably evolved over the period between 2008 and 2020, with associated trends potentially accounting for a good portion of the changes in measurements observed. Finally, it should be noted that the historical archiving of hyperlinks is a relatively recent practice. The oldest hyperlink found by Science-Metrix for the online outputs examined dated from 2015, and the quasi majority of these links were recorded in 2018 or later (they may well however have existed prior to that date). It should be noted that web citations are often performed by project team members legitimately diffusing links to their own productions and findings. Note that other methodological limitations to this analysis are noted in the methods section below. Given all these limitations, Science-Metrix has, much like for altmetrics, computed indicators revolving around those records that receive at least one or more web citation.

First examining web citations to project websites, 30 out of 75 CRA 2012-2016 websites had been the target of at least one hyperlink (40%). A much greater share of BiodivERsA project websites had received at least one web citation (83%). Looking at mean and median web citations levels of those websites that had received at least one hyperlink, the Forum and BiodivERsA project websites appeared to perform roughly at similar levels. BiodivERsA did somewhat better on the mean (13.4 versus 10.2 for the the Forum) but slightly worse on the median (8.5 versus 9.0 for the Forum websites). The low share of the Forum project websites to have received a web citation point towards a target for future improvement.

Table VII Belmont Forum achievements in fostering online attention towards non-journal outputs, 2012–2020

Type	Funder	Online count	Linked 1+	Share 1+	1+ mean	1+ median
Project websites	BF	75	30	40%	10.2	9.0
	BD	12	10	83%	13.4	8.5
Media	BF	91	34	37%	2.8	1.5
	BD	19	1	5%	2.0	2.0
News outputs	BF	43	24	56%	9.3	5.0
	BD	4	1	25%	3.0	3.0
Policy reports	BF	17	7	41%	20.7	2.0
	BD	11	4	36%	13.0	3.5
Research tools	BF	12	8	67%	3.5	3.0
	BD	8	4	50%	31.5	14.5

Notes: Online count is the count of outputs in the category that were made available online. Because not all outputs have been made available online, this count is often lower than the total count previously reported for that output category. Share 1+: share of output counts hyperlinked at least once. 1+ mean and 1+ median: mean and median of hyperlinks received within the subset of outputs with at least one hyperlink. Hyperlink counts are subject to multiple limitations, please see the methods section.

Source: Prepared by Science-Metrix using the BFGO database, Ubersuggest and web queries

Moving to other research outputs made available online, the Forum outputs fared better across the four categories considered here. Looking at shares of outputs to have received at least one web citation, the Forum outputs were ahead of BiodivERsA on this dimension in all four main categories of work considered. The Forum's lead was clearest for social media and video outputs, with a share of 37% of such works cited at least once online, against 5% for BiodivERsA. For journalistic news outputs, the Forum content was hyperlinked in 56% of cases, compared to 25% for BiodivERsA. Shares of content hyperlinked once or more was 67% (the Forum) versus 50% (BiodivERsA) for research tools, and 41% (the Forum) versus 36% (BiodivERsA) for policy reports.

Turning to mean and median levels of web citation within the subset of outputs that has received at least once hyperlink, closer jockeying was found between the two funders. BiodivERsA was clearer ahead for web citation levels towards research tools, with a mean of 31.5 (against 3.5 for the Forum) and a median of 14.5 (against 3.0 for the Forum). The Forum was ahead for its journalistic news outputs, with a mean of 9.3 (against 3.0 for BiodivERsA) and a median of 5.0 (against 3.0 for BiodivERsA). The Forum recorded higher observations for mean level of citations in the categories of social media (2.8 to 2.0) and policy report outputs (20.7 versus 13.0), but was behind on the median in the same categories (1.5 to 2.0 on social media; 2.0 to 3.5 on policy reports). This discrepancy indicated that hyperlinks to the Forum outputs in these two categories were skewed towards a few items, whereas they were slightly more evenly distributed in the case of BiodivERsA outputs.

Overall, the picture that emerges from this web citation analysis is that the dissemination outputs of supported by the Forum publications were effective in finding an online audience, while project websites were not so. It must be remembered that the type of web citations (hyperlinks) researched here amounts to just one particular modality of uptake and reception that can be achieved by online research outputs, let alone research outputs more broadly defined.

Case studies of web impact, stakeholder and outcomes achievements

In terms of web impact, the IPBES 2018 *Assessment Report on Land Degradation and Restoration*, to which DEVIL member Robert Scholes contributed as an editor and author, received a considerable amount of online attention, the most seen as part of this study. However, the writing of this report was a highly collaborative endeavour, and so Science-Metrix has not considered it to be an adequate example of the specific contributions made by the Forum-funded initiatives. Next behind the IPBES report, Pan-Arctic Options' investigator Paul Berkman's article published in the *Science & Diplomacy* trade journal has also seen considerable attention, including from the Wikipedia page on the concept of the 'Arctic'. Most of the online attention recorded can be traced back to the original Wikipedia entry, which has been reproduced in entries on wiki websites that derive their content from Wikipedia. An opinion piece published by Berkman in the *Alaska Dispatch News* was also moderately hyperlinked (5 hyperlinks).

The Deltas project produced a YouTube video, 'Why Do Rivers Have Deltas?', that received 25 hyperlinks, including from the Hebrew-language entry for 'deltas' on Wikipedia. A *Nature Climate Change* publication by the project team also attracted attention in journalistic outlets after an initial write-up by *Scientific American*. The *Scientific American* piece was hyperlinked 17 times, although online attention fuelled in part by a controversy surrounding the tone of the coverage employed by a *New York Times* editorialist.

Turning to BFGO' record of outcomes and stakeholders engaged by projects supported by the Forum, Science-Metrix notes that members of the Gold Matters project have exchanged with

policymakers and local regulators in Burkina Faso and Brazil. They have notably alerted these audiences as to the contribution of the mining industries to insecurity in the first country. In Uganda, they have assisted a women-led NGO in devising mining practices to reduce mercury exposure. Finally, they have contributed to increasing supra-national coordination capacity on the topic of extractive industries in organizations including the African Development Bank, the World Bank or the United Nations Environment Programme.

Members of the Urbanising in Place project (Nexus2016 call) have contributed to the launch of Citizen Soil Clinic network in London. This initiative will serve as a platform for soil-oriented citizen science database. These team members have also engaged municipal stakeholders (Greater London Authority) as well as conservation-oriented charitable organizations. Project members in Argentina and Belgium also interact with local municipal actors, in Rosario and Brussels respectively. Project members have been active in diffusing the outcomes from their interactions and workshops through journalistic interviews. Their proposal for the formation of a new Centre for Agroecology in Brussels has seen some amount of online attention (8 hyperlinks).

The In-Source project has contributed to energy transition and water management practices in New York City, Vienna and different locations in Germany. In New York, team members have supported the municipal Department of Environmental Protection and the local chapter of the American Institute of Architects in planning for rising waters and for wastewater management. A New York Institute of Technology news release on the project attracted the attention (e.g. hyperlinks) of publications such as a *Scientific American* blog, which was in turn hyperlinked 20 times, by sources including *American Infrastructure* magazine.

The NILE-Nexus project (Mountains2015) has modelled the Blue Nile's river flow, providing findings of direct relevance to at least 4 Ethiopian public sector partners, including the Ministry of Waters Resources, Irrigation and Energy. A blog post with contribution from team members was hyperlinked 6 times, including by Ethiopian news outlets.

Summary on the effectiveness and impact of wider dissemination activities

Considering shares of the Forum publications available in an OA modality, as well as mentions to these publications, the Forum appeared to have been effective in disseminating knowledge and fostering discussion of supported research beyond academic circles. Its performances could nevertheless improve, as shown by the few instances where it was surpassed in each category considered.

Turning to online uptake of outputs oriented towards dissemination and participatory or local engagement, the main conclusion was that effectiveness and impact appear to have been achieved, but that definitive evidence of these achievements could not yet be collected. Future substantiation of the findings reported here and improved monitoring of these dimensions are therefore crucial goals in the near-future for the Forum.

Comparison with the BiodivERsA 2015 review[59] highlighted the cursory character of many answers provided by the Forum teams on stakeholders and outcomes in their mid-term or final reports. More stakeholders and outcomes per project are reported about in the Forum reports than in the BiodivERsA review, but less information is available about the Forum achievements. Notably, the Forum self-reported entries provide only restricted information on how stakeholders are engaged and outcomes are achieved. Additionally, one would expect outcomes to be achieved in partnership with stakeholders, but these connections are seldom

[59] Lemaitre F. & Le Roux X. (2015) Analysis of the outputs of BiodivERsA funded projects: BiodivERsA 2008 joint call on "Biodiversity: linking scientific advancement to policy and practice". BiodivERsA report, 63 pp.

clarified in the Forum self-reported entries. Comparison with the BiodivERsA review showed that this funder has done possibly a better job at tracking examples of co-design with stakeholders and policymakers in funded project. It cannot be determined, however, whether this difference was due to discrepancies in reporting practices or in project practices as such. Nevertheless, it also remained that the sheer quantity of outcomes and stakeholders recorded through project reports is impressive and point towards likely achievements in effectiveness.

As a way to circumvent some of these shortcomings, "web citations" (hyperlinks) to the Forum-funded online outputs were also tracked, with the hope that they would capture instances of uptake in a broad audience of potential users. The BiodivERsA review provided a partial benchmark against which to appraise these achievements, enabling to identify outputs that were then evaluated for the web citation impact. On all categories except project websites, the Forum online outputs had seen a higher share of their numbers being referred through at least one hyperlink. The Forum online outputs had a higher mean number of such hyperlinks (within the subset of outputs cited at least once), although tended to have a lower median, meaning hyperlinks were more skewed towards a few units. These findings also supported a cautious conclusion of effectiveness.

EQ6: Policy effects and outcomes: To what extent did results of the Belmont Forum foster policy debate or developments at international and national level or facilitate policymaking/implementation?

Key section findings

A great proportion of scientific publications supported by the Forum were cited in policy documents, indicating that the research had been useful input for decision-making and argumentation by a range of governmental agencies, IGOs and think tanks. The share of articles cited by at least one policy document was close to a third for those publications from CRA 2012-2014 projects. Notable policy citations originated from the EU; the FAO; the UNEP; and the World Bank.

In more pointed comparisons, the share of ecology papers cited in policy documents for the Forum (36.8%) exceeded that for BiodivERsA (29.6%) but the difference was not statistically significant. The Forum papers were also cited to a similar extent as ERC publications (34.2% to 33.1%) in a selection of three subfields, and recorded higher levels than for NSF and NERC papers (18.6% and 20.6%, respectively, resulting in a statistically significant lead for the Forum).The data suggest that the collaborative models implemented by the Forum and BiodivERsA may be playing a positive role in the extent of policy uptake, just as the model of ERC focusing on research excellence might be. Further research would be needed to confirm the relative influence of these programmes' characteristics on policy uptake. Science-Metrix is currently conducting such work using econometric modelling to trace the impact of cross-disciplinary research, and other factors, on policy uptake. Our findings suggest a positive effect of scientific impact and cross-disciplinarity, computed at the paper level, on the subsequent uptake of research findings in decision making (in preparation).

Benefitting from the recent emergence of databases that systematically capture and parse policy documents (including white papers, parliamentary and other institutional deliberation transcripts; and legislative texts) made available online, it was possible to examine policy citations made towards journal publications supported by the Forum. Policymakers commonly refer to evidence from scientific publications in supporting their arguments, using references that mirror those found in journal articles and that can now be recorded on a large scale. The Overton database used for this component of the evaluation systematically indexes more than

two million documents produced by governmental agencies of all levels; e.g., think tanks, IGOs, some NGOs, and others. It does display some level of bias towards documents written in English and/or originating from Anglo-Saxon countries, although many European countries as well as Japan are also well represented in the database. While it does capture some amount of executive and legislative activity, a large portion of the policy documents it contains are grey literature reports that aim to synthesize research findings for policymakers.

The analyses conducted on policy citations (*Table VIII*) were designed with a different approach to those presented so far. Policy citation data was retrieved for the whole of the Forum publications (minus a small number of articles for which matching to Overton was not possible). In addition, a few comparators were chosen to construct small data sets for which policy citations could be retrieved and that could, in as much as was feasible, matched to comparable subsets of the Forum publications in terms of distribution by subfield and publication year. BiodivERsA and the ERC were retained as the main comparators of interest for the Forum. NERC and NSF were also included as other important comparators that would be unaffected (or in fact perhaps even positively affected by) coverage biases in Overton. In other words, NERC and NSF would provide useful comparators while simultaneously providing insight into the extent to which coverage biases in Overton should be factored in when interpreting these findings.

Table VIII Share of Belmont Forum publications' (and of comparators' papers) cited at least once in policy documents, 2009-2019.

Group	Total N pubs	N pubs pol. cited	Share pol. cited	Stability intervals	p-value (BF lead)	Notable pol. citations
Descriptive statistics						
Belmont Forum overall	654	145	22.2%			
Belmont Forum, CRAs between 2012–2014	368	117	31.8%			EU; FAO; UNEP; World Bank
Comparison A: Ecology subfield						
Belmont Forum, CRAs between 2012–2014	68	25	36.8%	(25.3%, 48.2%)		as above
BiodivERsA, 2008 call	142	42	29.6%	(22.1%, 37.1%)	0.148	IPBES; Naturvårdsverket
Comparison B: Three main subfields						
Belmont Forum, CRAs between 2012–2014	199	68	34.2%	(27.6%, 40.8%)		as above
European Research Council	199	66	33.1%	(26.6%, 39.7%)	0.412	Arctic Council; EU; UNEP
National Science Foundation, US	199	37	18.6%	(13.2%, 24.0%)	0.000	FAO; UNEP; WMO
Natural Environment Research Council, UK	199	41	20.6%	(15.0%, 26.2%)	0.001	IUCN; UNEP

Note: Main subfields for comparison B: Ecology; Environmental Sciences; Meteorology & Atmospheric Sciences. The number of papers used to compute the share of papers cited in policy documents is lower than the total number of papers for each comparator since papers with no DOI in Scopus could not be queried in Overton (the policy database). *P*-values are based on a one-tail test for a null hypothesis of the Forum smaller of equal to comparator.
Source: Prepared by Science-Metrix using the Overton and Scopus (Elsevier) databases

As *Table VIII* shows, 31.8% of publications supported by the Forum resulting from the 2012-2014 CRAs were cited at least once by a policy document from the Overton database. Notable policy users of research supported by the Forum include the European Union, the Food and Agriculture Organization of the United Nations, the UN Environmental Programme, and the World Bank. Looking at the full publication set supported by the Forum, the policy citation rate drops to 22.2%. However, it should be noted that it may take three to four years, or even longer, for a peak in the share of papers cited in policy documents to accrue after the publication date. Therefore, publications supported by the Forum from recent CRAs account for this decrease because they had not accumulated citations over a long enough period at the time of writing these lines.

A first comparison of policy citation rates was performed against BiodivERsA's accomplishments on this dimension, restricting the analysis to publications in the subfield of Ecology; most of BiodivERsA's papers were in this area (comparison A in

Table *VIII*). Within this analytical subset, the Forum CRA 2012-2014 publications reached a policy citation rate of 36.8%. BiodivERsA papers were cited in a proportion of 29.6%. The Forum publications' lead in this comparison was likely, although not statistically significant (*p*-value of 0.148) and therefore not free of uncertainty. Given restriction in available data, we consider it is safe to conclude that both funders sharing a similar model of funding international and cross-disciplinary collaboration lead to similar levels of policy uptake.[60] When looking at comparison B in

Table *VIII*, we see that The Forum scores only slightly lower than it does in comparison A. Since it scored markedly above NSF and NERC in this case, one might hypothesize that the collaborative model of the Forum and BiodivERsA favour the policy uptake of research findings relative to more general funding models (ERC in is a separate category being focused on research excellence). Notable sources of policy citations to the BiodivERsA publications included the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Naturvårdsverket (Swedish Environmental Protection Agency).

In a second comparison (comparison B in *Table VIII*), the Forum publications' policy citation achievements were appraised in relation to those of the ERC, the NSF and the NERC. This comparison was restricted to the Meteorology and Atmospheric Sciences, Environmental Sciences, and Ecology subfields. The share of publications cited in policy documents was 34.2% for the Forum, 33.1% for the ERC, 18.6% for the NSF, and 20.6% for the NERC. The small lead of the Forum accomplishments to those of the ERC was not statistically significant (p-value of 0.412). The Forum's leads on the NSF and NERC observations, however, are large enough to leave no place for uncertainty (*p*-values of 0 and 0.001 respectively). The three comparators mobilised here received notable policy citations by UNEP; the Arctic Council (ERC); the World Meteorological Organization (NSF) or the International Union for Conservation of Nature (NERC).

Taken together, the findings presented in this section indicated that the Forum peer-reviewed publications reached high level of impacts within regulatory science and scientific advice circles. Overton data on citations from policy documents does not provide insight solely into legislative or executive use of research findings, but also include a large swath of grey literature. Therefore, these findings could not be interpreted as indications of direct impact on policy formulation and implementation, but rather of input provided into the first steps in the process of knowledge transfer.

EQ 15: What is the added value of the Belmont Forum (compared to other initiatives at various governance levels)?

Given the comparative approach already deployed in computing bibliometric indicators of effectiveness and impact, the findings previously reported when assessing effectiveness and impact of the Forum-funded projects are also of relevance in assessing their collective added value. Accordingly, Table IX recapitulates select findings from the analyses presented above. To emphasize added value, it presents the differences in observations between the Forum

[60] Pinheiro, H.N., Vignola-Gagné, E. & Campbell, D. (2020) *Using Overton policy citations in assessing the uptake of cross-disciplinary research in decision making*. Science-Metrix. In preparation.

publications and those of selected comparator groups instead of the empirical measurements recorded (and already presented above).

Table IX shows the Forum performances to be above those of comparators on most dimensions of collaboration examined. The Forum lead was clearest on North-South ICR (8 percentage points above parallel publications by the Forum investigators and 17 above prior publications, to take only the comparison groups most likely to engage in such collaborations). Lead was also very clear on inter-disciplinarity and multi-disciplinarity, with no lead below 6 percentage points. Here it must be considered that initiatives such as BiodivERsA and the ERC do have an explicit mandate to support these research practices. When considering the inclusion of authors from "non-research sectors' in peer-reviewed publications, the Forum publications had reached similar levels of diversity as concurrent publications by the Forum awardees (0.3 percentage points for papers written as a co-publication between any "research sector" and any "non-research sector" – ISR (Acad.LRC.ORC.Gov) & (NGO.IGO.Priv)) and had a clear lead on BiodivERsA papers (4 percentage points). The Forum publications were less likely to be written as collaboration with a private sector actor then for publications from these two comparator groups, but they were much more likely to be written together with an author based at an NGO (or at an IGO, comparatively to BiodivERsA).

In terms of uptake of the Forum findings within the scientific community, the Forum publications offered their best performance for this dimension on the $HCP_{10\%}$ indicator, surpassing even ERC publications (almost 4 percentage points above). On the CDI, which offers a balanced index of citation impact profiles across the full spectrum of publications for a given analytical group, the Forum publications did fall behind those of the ERC and BiodivERsA (7 and 1 percentage points below, respectively), but remained above many other comparator groups.

Table IX Added-value of scientific and societal outcomes Belmont Forum-supported projects, 2007-2020

Dimension	Non-BF publications by BF awardees		BF awardees prior publications		BiodivERsA, 2008 call		ERC		NSF		NERC	
North-South ICR	+8.3 p.p.	*	+16.5 p.p.	*	+34.1 p.p.	*	+14.8 p.p.	*	+26.3 p.p.	*	+22.6 p.p.	*
$HIP_{10\%}$	+10.2 p.p.	*	+10.1 p.p.	*	+11.6 p.p.	*	+9.4 p.p.	*	+12.7 p.p.	*	+10.5 p.p.	*
$HMP_{10\%}$	+6.3 p.p.	*	+7.4 p.p.	*	+7.5 p.p.	*	+14.1 p.p.	*	+12.1 p.p.	*	+8.5 p.p.	*
ISR (Acad.LRC.ORC.Gov) & (NGO.IGO.Priv)	+0.3 p.p.		+4.8 p.p.	*	+4.4 p.p.	†						
ISR (Acad.LRC.Gov.ORC) & NGO	+4.4 p.p.	*	+5.9 p.p.	*	+5.2 p.p.	*						
ISR (Acad.LRC.Gov.ORC) & IGO	0.0 p.p.		+2.2 p.p.	†	+4.2 p.p.	*						
ISR (Acad.LRC.Gov.ORC) & Priv	-4.2 p.p.		-1.5 p.p.		-2.7 p.p.							
ISR (Acad.LRC) & (ORC.Gov) & (IGO.NGO.Priv)	-1.0 p.p.		+1.3 p.p.		+2.3 p.p.	†						
ISR (Acad.LRC.ORC.Gov) & NGO & IGO & Priv	-0.1 p.p.		+0.1 p.p.		+0.3 p.p.	‡						
$HCP_{10\%}$	+15.7 p.p.	*	+12.9 p.p.	*	+10.3 p.p.	†	+3.6 p.p.		+11.9 p.p.	*	+11.0 p.p.	*
CDI	+6.8	†	+6.6	†	-1.1		-6.9		+2.8		+2.5	
OA%	+6.5 p.p.	*	+5.5 p.p.	†	-10.0 p.p.		-8.1 p.p.		+7.8 p.p.	*	-18.6 p.p.	
News $HAP_{10\%}$	+10.3 p.p.	*	+13.5 p.p.	*	+11.5 p.p.	*	+3.8 p.p.	‡	+4.7 p.p.	†	+3.4 p.p.	‡
Wikipedia $HAP_{10\%}$	+1.4 p.p.	‡	+1.6 p.p.	‡	-3.3 p.p.		-0.8 p.p.		+0.4 p.p.		-0.4 p.p.	
Share of policy cited					+7.2 p.p.		+1.0 p.p.		+15.6 p.p.	*	+13.6 p.p.	*

Note: Indicators in this recapitulative table are subject to limitations, please consult prior tables for more details. Shares of publications with a least one author with an affiliation in the category or categories of interest. Do note that authors may have multiple affiliations falling within multiple sectors. Acad: universities and academic institutions. Gov: government research centers (mostly) and governmental agencies. LRC: Large research centers, often government-funded and quasi-academic, such as the Max-Planck network of institutes. ORC: Other research centers, whose institutional status could not be clearly established but which appeared to be independent research centers; charity-based; or (less often) government-funded. IGO: inter-governmental organisation. NGO: non-governmental organisation. Priv: private sector. One-tail test for differences of means between Belmont Forum and each comparator: * p<0.01, † p<0.05, ‡ p<0.1

Source: Prepared by Science-Metrix using the Overton; PlumX (Elsevier); and Scopus (Elsevier) databases

Turning finally to indicators of broad dissemination and societal uptake of findings, the Forum publications displayed a high level of impact within regulatory science and governmental research documents. The Forum was trailed closely only by the ERC (one percentage point behind, a non-significant lead for the Forum). A lead to BiodivERsA (7 percentage points) cannot be considered as statistically definitive but nevertheless was meaningful enough to comfort an observation of added value for the Forum publications. The share of the Forum peer-reviewed publications to be amongst the most mentioned in journalistic news items for their subfield (news $HAP_{10\%}$) was consistently higher than for comparators (3 percentage points or higher). The share of the Forum publications to be mentioned in Wikipedia articles was roughly equivalent for the groups included in the comparison, although they were below articles by BiodivERsA on this dimension (3 percentage points behind). The Forum publications show arguably their weak spot in terms of OA availability. Shares of the Forum publications available under OA modalities were 8 percentage points below the corresponding share of ERC papers, and almost 19 percentage points below that of NERC articles.

In short, it can be concluded that the Forum support offered clear added value within the global landscape of science funding and a unique combination of project-level strengths and achievements. The Forum funding has managed to simultaneously foster: research excellence of the highest order; North-South collaboration; multi- and inter-disciplinarity in the production of formal knowledge outputs; collaborative work most notably with NGO; and strong uptake in regulatory science and research for policymaking. The few areas for improvement include availability of peer-reviewed publications in OA and some aspects of online knowledge dissemination.

As was the case for effectiveness and impact, more definitive evidence on non-research outcomes of the Forum projects are still needed at this point. Options available for the Forum to obtain more reliable data on non-research outcomes can be envisioned as follows:

- Foster more quality self-reporting by funded teams in their project reports;
- Expand the Forum Secretariat capacities and resources to enable it to actively monitor non-research outcomes itself, as is being done in comparable initiatives such as Future Earth or BiodivERsA;
- Commission a follow-up external evaluation that will focus solely on this dimension. Such an evaluation would most certainly need to match the current evaluation in scope and resources engaged to achieve its goals.

Appendix – Methods

Analytical design

Comparative strategies

Access to baseline and benchmark data sets greatly increases the interpretative value of performance measurements performed as part of programme evaluations.

Science-Metrix has examined the structure, features, and data available on programmes that could act as comparators to the the Forum's CRAs for the purpose of the current evaluation. We have considered Future Earth, FP7, and H2020 projects on climate change; NORFACE; and the BiodivERsA programme for use as an evaluation comparator. Science-Metrix considers that the BiodivERsA programme, funded through the European Commission's Framework Programmes (i.e., FP6, FP7, and H2020), would act as the best comparator from this group, for the following reasons:

- the duration of its awards;
- the monetary value of its awards;
- the use of a similar joint call structure that combines funders from multiple countries;
- the typical composition of supported teams (i.e., interdisciplinary and international); and
- the thematic focus of its awards.

Future Earth was found to support projects of a larger scale and much longer duration than the Forum's CRAs. Norface's focus on the social sciences does not make it directly comparable. A well-made selection of thematically relevant FP7 and H2020 projects might be the other interesting option for building a comparator, although this comparator will amount to an analytical construct rather than a unitary programme like the the Forum.

Primary comparator using bibliometric and altmetric indicators on peer-reviewed publications: In recommending the use of the BiodivERsA programme as a primary comparator for the the Forum's CRAs, Science-Metrix made note of the availability of an in-depth review of outputs from the 2008–2009 joint call of that programme, published in 2015. This review contains comprehensive lists of project outputs, including publications and other forms of research by-products, which were, as for the Forum, desirable given the call's focus on linking scientific advancement to policy and practice. The availability of such curated and vetted lists of outputs greatly enhanced the robustness of the comparative analysis. Although the reference period for the review of the BiodivERsA call (i.e., 2008–2009) is prior to any of the Forum's CRA calls, it should be kept in mind that many of the standard bibliometric indicators computed by Science-Metrix are normalised by year and field of science, with the world as a reference. Therefore, the bibliometric procedure intrinsically controlled for any temporal trends that might have otherwise differentially affected the findings for BiodivERsA and the Forum's CRAs.

Primary comparator using custom altmetrics on non-article outputs: For the specific case studies performed using custom altmetrics on non-article outputs, the comparison of the 2008–2009 call of BiodivERsA to the 2012 CRAs was ideal since information on these types of outputs was, in both cases, gathered roughly seven to eight years following the respective call (i.e., in 2015 for BiodivERsA and in 2020 for the CRA). Of course, the lag to uptake was different. One caveat for the custom altmetrics was that the lag in measuring uptake will be longer for BiodivERsA than for the CRAs. However, the custom altmetrics aimed to provide a qualitative understanding of longer-term uptake rather than a fool-proof quantitative benchmarking

against a comparator. Also, the 2015 BiodivERsA review has appraised non-articles outputs in a narrative mode rather than a quantitative one. Data from the review could therefore only be used as a reference in the qualitative portion of the analysis. BiodivERsA project websites were nonetheless examined for information on project non-journal outputs, but multiple websites had closed since the completion of the research projects. Comparisons were only made with caution in this subset of analyses.

Longitudinal analyses and within-group comparisons: Some comparisons aimed to measure the extent to which plans for transdisciplinarity and collaboration, as they appear in grant applications, have materialised in peer-reviewed scientific outputs. In those cases, publications of the investigators supported by the Forum were retrieved for the five years period prior to the start of their award. Comparison were conducted on achievements during and prior to the award period.

Comparisons between pre- and award periods do not control for broader trends. For example, an increase in the international co-publication rate of the supported researchers could have occurred in the absence of the Forum since this is generally a clear pattern at world level. Also, it was possible that an effect was not observed because the Forum supported individuals who were already exhibiting a strong propensity towards international cooperation.

To help control for such confounding factors, a comparator group was also elaborated from parallel publications by the Forum investigators published during the award period but with no mention of the Forum funding in their acknowledgements. This control group's differences in performances were compared against the Forum pre- period as can the publications supported by the Forum. In effect, this led to a difference-in-difference analysis that controls for some degree of local and global trends when comparing temporal changes between non-Forum and the Forum publications by supported investigators. The group of non-Forum publications acted as a counterfactual controlling for differing characteristics of individual awardees. It can be noted that this counterfactual was not perfect, given that the Scopus data on funding acknowledgement displays known recall issues. There is a possibility that non-Forum publications were in fact the Forum publications because Scopus coding of funding sources has been imperfect in those cases; or that investigators themselves failed to appropriately mention Belmont support in relevant papers' acknowledgements. Manual spot checks in the non-Forum publications records revealed that such cases are unlikely to be widespread. Conservative interpretations of findings on the Forum publications vs non-Forum publications by awardee investigators would account for all scenarios in this respect.

Secondary comparators using bibliometric and altmetric indicators on peer-reviewed publications: Additional comparators to the CRAs for the peer-reviewed scientific outputs were identified through a data set approach. This approach entailed delineating a global set of publications with similar topics to those of CRA publications (falling mostly in the Science-Metrix categories of Ecology; Environmental Sciences; and Meteorology & Atmospheric Sciences). These publications were retrieved from Scopus using keyword-based queries and citation relationships. Science-Metrix was able to identify other notable funders within this topical data set using the information from the publications' acknowledgements in Scopus. It should be noted that these comparators were not selected according to their similarity (organisation, mission, or topical focus) to the Forum model, but simply based on their occurrence within the delineated publication set. Finally, the thematic set of comparable publications allows the calculation of world reference performances level that provides a baseline for benchmarking.

Analytical periods

Primary and data set comparators using bibliometric and altmetric indicators on peer-reviewed publications. The outcomes and impacts captured through bibliometric and altmetric methods are realised after the start of research projects, sometimes many years afterwards (often after the end of grant support). These lags between support periods and outcome realisation periods are captured in Table X, showing that even projects in the earliest CRAs (Coastal 2012 and Freshwater 2012) were still producing journal publications and that they reached volume peaks in 2016 or 2017. Therefore, robust bibliometric, altmetric, and network analysis findings were only produced with careful consideration is afforded to the influence of analytical periods on the performance levels captured. The next paragraph summarises Science-Metrix's proposals concerning the selection of periods in conducting benchmarking analyses, while the paragraphs that follow provide more detail into the guidelines and observations that support these design choices. The sub-sections that follow present the analytical periods retained for comparisons using custom altmetrics and non-comparative analyses.

Science-Metrix conducted its bibliometric, altmetric, and network analyses on the outputs of those projects funded by CRAs issued in 2012, 2013, and 2014 (amounting to five CRAs). For citation- and altmetric-based indicators, only publications from those five CRAs published between 2013 and 2016 were retained, to allow for sufficient outcome realisation periods. For bibliometric indicators that rely solely on publication-level data, all publications from the five CRAs issued between 2013 and 2018 were appropriate for inclusion.

For BiodivERsA, publications from the 2008 call and made available between 2009 and 2019 were included (or to 2016 for citation indicators). For all other comparator groups, publications between 2014 and 2019 were used (or to 2016 for citation indicators).

The following paragraphs provide further details on the justifications for these design strategies.

For publication-based indicators, the main challenge is to establish quasi-causal links between publications and funding, given how investigators typically hold multiple concurrent grants and how these may support multiple experiments and activities that may not fall neatly within the boundaries of one project or another. BFGO records of publications are clearly linked to the Forum support, but we cannot be as certain for records obtained by the enrichment strategies. For instance, Science-Metrix found multiple project websites containing mentions of publications from before formal award start. For most publications, it is expected that the production of experimental findings and their publication will take place after a lag from formal award start. Therefore, Science-Metrix only retained publications from a project issued in the year immediately following the year of CRA issuance. For instance, publications associated with the Coastal 2012 CRA were only included in the analysis if they were published in 2013 or after.

For citation-based indicators, outcome realisation periods are longer than for publication-based indicators. Citations only start accruing once publication of findings has been achieved, adding a second lag to the publication lag itself. Accordingly, at least four full years—or arguably even more—should have elapsed since the start of a support programme before assessing the citation profiles of relevant publications. Given this, only the projects awarded in the five 2012–2014 CRAs had enough publications (165) matching the above criteria for computing citation-based indicators using Scopus data. For more recent CRAs, only eight publications matched the above requirements. The final sets of journal publications used from the five CRAs in those years also varied depending on the indicator.

Publication of findings must also have been achieved for altmetric "citations" to publications to accrue, as is the case for citation-based indicators. Additionally, expertise and experience with altmetrics methods in both the scientific and evaluation communities are still being consolidated, and standards have yet to emerge to guide practice. Yet, since historical hyperlink data is skewed towards web citations made between 2018 and 2020, it was appropriate there to include online content from projects funded by CRAs issue between 2012 and 2016.

Table X Journal-based publication outputs by CRA and year

CRA	2013	2014	2015	2016	2017	2018	2019	Total
Coastal2012	1	9	24	30	24	16	9	113
Freshwater2012	1	3	13	21	22	8	8	76
Food2013			7	22	38	48	12	127
Biodiversity2014		3	3	4	9	7		26
Arctic2014		1	5	18	20	15	9	68
Climate2015				3	50	72	30	155
Mountains2015			3	2	5	19	12	41
Nexus2016						14	8	22
T2S2016					2	4		6
Biodiversity2017					1		2	3
Not Identified *				3	5	10	4	22
Total	2	16	55	103	176	213	94	659

Note: "Not identified" refers to part of the publications retrieved from Scopus acknowledging the Belmont Forum but for which no clear criteria were found to assign them to one of Belmont's funded projects. Note that 2019 publications may not yet be fully indexed in Scopus as of January 2020.
Source: BFGO, Scopus, and Science-Metrix web queries

Primary comparator using custom altmetrics on non-article outputs Comparing the outputs from the BiodivERsA 2008–2009 call to the 2012 CRAs did, to some degree, advantage the former over the latter. This is because the outputs from 2008 grants likely accumulated altmetric citations over a longer time frame than those of 2012 grants. Additionally, since web citations are not associated with a time stamp other than the year of extraction, which will be the same (i.e., 2020) for all projects investigated through the custom altmetrics, it was not being possible to control for the effect of this lag. Furthermore, a large reference data set is not available to control for time effects in the custom altmetrics analysis. Accordingly, the custom altmetrics analysis aimed to provide a qualitative understanding of the longer-term uptake of non-article outputs rather than a fool-proof quantitative benchmarking against a comparator. Here again, the evaluation's other lines of evidence will come in handy to interpret the findings.

Statistical robustness and mitigation of risks

Science-Metrix has found that publication volumes in some analytical subgroups (specific CRAs or years) are sometimes small, possibly compromising robustness of findings in some cases. Science-Metrix will typically present robust stability intervals (analogous to 95% confidence intervals but constructed using bootstrapping) for the computed indicators, helping to obtain more robust interpretations of results associated with smaller volumes of publication. In some cases, Science-Metrix may decide not to compute certain indicators if publication volumes are too small. For example, we never compute citation-based indicators for publication volumes of less than 30, due to the undue influence of outliers on averages in such samples. In

performing comparisons to the primary comparator, for bibliometric and altmetric indicators on peer-reviewed publications, some statistical comparisons were also reported using robust *p*-values and effect sizes obtained using bootstrapping.

Our bootstrapping procedure: The application of standard statistical tests of significance and effect sizes to the large-scale descriptive findings produced by bibliometric exercises is not straightforward and has been the subject of debates in the expert community. Following an increasingly broadly used approach, Science-Metrix has recently developed methods that can be applied to bibliometric data sets and findings so as to measure the degree of uncertainty they are to be associated with.

For citation-based indicators, Science-Metrix advocates the use of stability intervals to assess uncertainties. Stability intervals inform on the uncertainty of bibliometric indicators by providing a range within which a computed score could likely fluctuate in response to a change in the underlying set of publications that was used to compute it. Stability intervals are built by randomly resampling, with replacement, a group's papers to produce many resamples (e.g., N = 1000) of equal size to the group's number of papers. The various indicators to be produced are then computed for each resample to produce an empirical distribution of the scores. This enables the computation of a 95% stability interval—that is, the interval containing 95% of the resamples' scores.

For example, a stability interval could be used to build the range of scores including 95% of the likely values for the proportion of highly cited publications published by two groupings of projects supported by the Forum (say, CRAs). If the share of highly cited papers in CRA A equals 13% with a 95% stability interval ranging from 11% to 15%, and the share of CRA B equals 17% with an interval ranging from 16% to 18%, then it would be safe to conclude that CRA B would perform better than CRA A even if the underlying sets of publications were to change. The rule of thumb is that if the 95% intervals of the groups being compared do not overlap, then the observed difference is highly likely to remain visible should the underlying data be altered. Because they are built empirically, stability intervals do not rely on the assumptions that the study samples are random and follow a specific distribution. However, they assume that the observed data are representative of the larger populations to which they belong.

The approach outlined above can also be applied to altmetrics computed for journal publications. Here, however, statistical testing will not be able to mitigate some of the intrinsic limitations that are widely associated with altmetrics. For instance, altmetric data sources are known to suffer multiple biases (e.g., notably in terms of language/country coverage) whose extent are still not fully documented. As a result, the application of altmetrics to formal evaluations remains rather limited. Accordingly, all quantitative findings stemming from the altmetric analyses to be performed in this review should be considered with due care. More explicitly, this implies that the quantitative findings coming out of the altmetric analyses will have to be substantiated through triangulation with this evaluation's other lines of evidence. Accordingly, quantitative results from the altmetrics component are best regarded as an experimental complement to the more established bibliometric indicators. Nevertheless, note that altmetric analyses will also generate a large volume of qualitative information (contextual information on the broader impacts of the CRA outputs) of high value to this review.

Data sources

Starting data set: BFGO database

The Forum, together with contractor LUX Consulting Group Inc., maintains the BFgo database, a repository of information on the grant applications and research projects supported by the

Forum. Extracting data from progress reports filled by grantees, BFgo contains a wealth of information on outputs produced by supported projects. This information is not fully comprehensive, however, as not all project teams have filled in the required reports. Additionally, progress reports for projects funded through the 2012 CRAs were also not recorded in BFgo. Nevertheless, BFgo provided the starting list of journal-based outputs matched to Science-Metrix's bibliometric database, as well as the starting list of non-journal outputs to be examined through the custom altmetric analyses.

Publication databases

Currently, only two databases offer an extensive coverage of international scientific literature and index all bibliographic information necessary to perform advanced analyses on scientific activity: the Web of Science, produced by Clarivate Analytics, and Scopus, produced by Elsevier.

Science-Metrix will use Scopus for this contract, based on its extended coverage, which in turn supports a more balanced representation of countries and disciplines. Scopus also includes some features that are helpful in the constitution of researchers' portfolios (author IDs, consistent links between authors and their affiliations, greater availability of full names rather than just initials). Scopus provides comprehensive coverage of the scientific literature: for the period 2012 to 2018, it contains close to 17 million publications, covering about 35,000 peer-reviewed journals in 176 disciplines.

Altmetrics databases

Science-Metrix will combine information from the PlumX and Overton databases when retrieving information on the altmetric impact of research supported by the Forum.

PlumX maintains a database recording the uptake of journal-based scientific outputs beyond the scientific literature in, for example, social media, blogs, news, and educational resources. These data, because they track usage beyond academic circles as traditionally captured in bibliometric indicators, are often referred to as alternative data (or altmetrics). Included in the database's coverage are platforms such Facebook and Twitter, a selection of blogging platforms, journalistic and news websites, Wikipedia, Reddit, Stack Exchange, and library holding databases. The PlumX database contains records for 52.6 million individual pieces of research output, which in total have been linked to over 9.4 billion altmetric captures. Of these research outputs, 83.2% have registered altmetric uptake for at least one of PlumX's metrics. PlumX metrics associated with Scopus records are integrated into Science-Metrix's implementation of the latter database.

The Overton database consists of more than 1.65 million policy documents. These policy documents include white papers from international multi-lateral organisations, as well as guidelines from city councils, parliamentary transcripts, and other classes of the so called "grey literature". Around half of these documents make citations to academic or scholarly publications. Just over 2 million distinct journal-based publications are cited by at least one policy document in the database. The Overton database does display bias towards English-language documents originating in Anglo-Saxon countries. However, as the only robust database of its kind currently available, its addition to the set of data sources used for this evaluation will lead to valuable insights into the knowledge transfer between academia and the policy-making world supported by the Forum.

Hyperlink analysis

The Ubersuggest portal is a commercial tool originally geared towards the production of web analytics for online strategic marketing purposes. It contains a version of the Google hyperlink crawl database and can therefore be used to retrieve records of web citations towards websites and other online content (YouTube videos, blog posts, or any content associated with a distinct URL). Use of this database as an evaluation or investigative tool is exploratory and subject to multiple limitations highlighted in the main text and below. Hyperlink archiving is a recent practice, and most hyperlinks retrieved in this study were recorded between 2018 and 2020, but the oldest one registered dated from 2015.

Data preparation

Constitution of the Belmont Forum journal-based publication set

In performing its preliminary exploration of CRA-supported projects/outputs, Science-Metrix appraised and processed data from the BFgo database to recast it in the structured form required for analysis. Science-Metrix also conducted systematic web searches on additional outputs by funded projects that were not mentioned in progress reports (and therefore were not inputted into BFgo). Through this enrichment strategy, Science-Metrix more than doubled the volume of peer-reviewed scientific publications to be included in the analysis and doubled the number of other forms of outputs (including outputs such as engagement workshops, social media posts, policy reports, or conference presentations). In short, given some of the limitations of the data sets made available to Science-Metrix, its analysts have completed the initial phase of data collection as input towards the finalisation of analytical strategies.

Science-Metrix identified 324 traditional scientific articles (published by peer-reviewed journals) as outputs of the CRA projects listed in the BFgo database. Of these, 289 could be matched to Science-Metrix's production version of the Scopus database (Elsevier), which will be used in computing bibliometric indicators.

Information on CRA outputs was only available for roughly half of the projects from the CRAs included in BFgo's output section. As a result, Science-Metrix enriched the initial data set by retrieving additional information on CRA-funded publications using other data sources. First, it examined project websites maintained by supported investigators. Then, it parsed online awards databases maintained by national and European funders (including the NSF Award Search portal and OpenAIRE) for records relevant to the Forum-funded projects. It systematically examined databases from all funders that are members of the Forum. Finally, it queried its Scopus database (see section E.7) to retrieve any records making mention of Belmont support (particularly in the acknowledgement sections of articles). Together these strategies identified another 413 records of publications to be added to the starting data set. Out of these 702 publication records, 659 can safely be attributed to the CRA projects and contain enough information to allow for full bibliometric and altmetric analysis. Of course, only a subset of these publications is amenable to citation analysis because some are too recently published to have accumulated citations.

This bibliometric appraisal of the scientific production of the Forum is based on a set of scientific publications retrieved from four main sources: BFgo (the Forum grant application and monitoring database), websites of research projects funded under the Forum, funding organizations websites, and the acknowledgment section of Scopus Database. The publications collected in these primary sources were matched to the Scopus database, resulting in a set of 658 peer reviewed that based the computation of bibliometric indicators of the Forum.

Constitution of the Belmont Forum non-journal outputs set

Apart from scientific publications in peer-reviewed journals, Science-Metrix found 607 records of other project outputs in BFgo. The enrichment strategy for these types of outputs relied more heavily on project websites since very few funders track the non-article outputs of their funded projects, and because there is no global database indexing these other forms of scientific outputs. Working from project websites, Science-Metrix coded 612 additional records of non-article outputs from CRA projects. It should be noted that project websites were found for 66 out of the 99 total projects that the Forum has supported between its 2012 and 2018 (SEI only) calls. It can be noted that the list of websites included ResearchGate project websites; BiodivERsA website project descriptions (for the joint 2017 call); or project descriptions included on host university portals.

BFgo also contains self-reported research outcomes from project investigators, and particularly lists of societal, economic, social or cultural outcomes; as well as lists of stakeholders engaged in the project. The level of attention afforded to project reporting varied greatly from one project to the other; and that respondents tended to focus on some of the output and outcomes types mentioned above at the expense of others. Finally, it might have been expected that descriptions of outcomes would mention stakeholders with which they would be co-produced, or demonstrate how collaborative activities have taken place. Such instances were rather limited, however.

Constitution of the comparative data sets

In addition to this dataset, the bibliometric indicators were calculated for a subset of publications restricted to the CRAs from 2012 and 2014 that accounts for more matured projects and, therefore, allowing more confidence in the results observed. If not otherwise specified, the Belmont publications mentioned here refer to this subset of publications (CRAs initiated between 2012 and 2014).

Comparator datasets were assembled to benchmark the indicators calculated for Belmont publications.

- Publications authored by the applicants of Belmont projects (principal investigators and co-applicants as they are listed in the Belmont website) after they have joined the program, but that were not supported under the Forum (with no mention of the Forum funding in the acknowledgements section).
- Publications authored by the Belmont's applicants in the five years preceding their entrance in the program.
- 426 publications funded under a similar funding program, the BiodivERsA 2008 call.
- A thematic dataset formed by 133,142 publications retrieved from Scopus database published after 2014. These publications were selected based on keywords used along their titles and abstracts that typically represent the description of the largest Belmont projects (in terms of number of publications retrieved for this study) and the abstract and titles of these same publications.
- Finally, a subset of this thematic dataset containing publications funded by the main 20 funding agencies (in number of funded papers), including three funders to be used as main reference points: the National Science Foundation, the European Commission (overall) and the European Research Council (as a subset of the EC).

Limitations in the analysis of citations from policy documents

Overton does display some level of bias towards documents written in English and/or originating from Anglo-Saxon countries, although many European countries as well as Japan

are also well represented in the database. While it does capture some amount of executive and legislative activity, a large portion of the policy documents it contains are grey literature reports that aim to synthesize research findings for policymakers.

The analyses conducted on policy citations were designed with a different approach to those presented elsewhere. Policy citation data was retrieved for the whole of the Forum-publications (minus a small number of articles for which matching to Overton was not possible). In addition, a few comparators were chosen to construct small data sets for which policy citations could be retrieved and that could, in as much as was feasible, matched to comparable subsets of the Forum-publications in terms of distribution by subfield and publication year. BiodivERsA and the ERC were retained as the main comparators of interest for the Forum. NERC and NSF were also included as other important comparators that would be unaffected (or in fact perhaps even positively affected by) coverage biases in Overton. In other words, NERC and NSF would provide useful comparators while simultaneously providing insight into the extent to which coverage biases in Overton should be factored in when interpreting these findings.

Indicator definitions

Publication output volume

Number of publications

This indicator shows the number of publications for a given entity, calculated using a method called full counting. Using this method, each country, economic sector, or research organisation that has a researcher on the list of authors for a given paper gets a full count (1 publication) for that paper. For example, if a paper is authored by two researchers with addresses in the United Kingdom, one from Spain, and one from the United States, the paper will be counted once for the United Kingdom, once for Spain, and once for the United States.

Share of open access publications

Open access (OA) as a topic in science policy has grown immensely in importance in recent years. Science-Metrix has an intense and long-standing interest in OA and is uniquely well placed to conduct bibliometric research projects on this topic. 1science, Science-Metrix's spin-off company, has constructed a database of peer-reviewed, OA publications using a web harvester to collect and characterise papers from the web. The definition being applied is a simple one: a publication is OA if it can be accessed for free and without any barrier, such as a subscription or registration.

Because the content of the 1science OA database is cross-referenced to the publications in the Scopus database, Science-Metrix is able to deploy its full arsenal of analysis tools to characterise the research enterprise as it evolves on both sides of the OA divide. For this study, Science-Metrix will retrieve the share (%) of the Forum journal publications that can be accessed through an OA licence.

Citation impact

All indicators of scientific impact used here are based on citations. An important assumption underlying such analyses is that citations are a good proxy for contributions to scientific knowledge. While it is true that citations are generally used to communicate the positive influence of one piece of research on another, citations are also sometimes used for other reasons. For example, one article may be contradicting another; the author would in that case use a citation to highlight the article being contradicted. Additionally, an article may cite many others, with some material constituting general background information and other material

constituting the principal foundation on which the new piece of knowledge is built. These varying citation behaviours are all treated equally in analyses of scientific impact, which are blind to the differences between them.

Scientific impact assessed on the basis of citations would therefore be better interpreted as contributions to and visibility within scientific discourse; it would not, for example, highlight a paper that is of good quality but that fails to get much visibility or recognition within the research community. In light of these considerations, the interpretation of scientific impact analyses should proceed with due caution.

Relative citation scores

Counting citations can be used as a proxy for measuring contributions to subsequent knowledge generation; however, because citation practices vary between the disciplines and sub-disciplines of science, simple counting would create unwanted biases in the results. To correct these potential distortions, individual publications are evaluated relative to the average citation rate for publications in the same subfield (using the Science-Metrix classification, or a custom taxonomy if tailored data sets are created) and published in the same year. This measure is known as the relative citation (RC) rate.

For all citation-based measures, a certain amount of time must be allowed for the published work to have an impact on subsequent research and for articles to be cited. A recent analysis conducted at Science-Metrix shows that only a small number of subfields reach citation peak within two years; that is to say, citation attention for papers is still continuing to increase even several years after publication, and therefore a measurement taken too early risks not effectively reflecting the total attention that a body of work will receive. For this reason, Science-Metrix will not compute impact statistics for papers published in 2017 or later, as they have not had sufficient time for citations to accrue.

Average of relative citations

The average of relative citations (ARC) is the average of the relative citation scores of all the articles published by a given entity. The ARC is normalised to 1, meaning that an ARC above 1 indicates that the entity's articles have higher-than-average impact, an ARC below 1 means that the entity's articles have lower-than-average impact, and an ARC near 1 means that the publications have near-average impact.

Because RC scores are known to be skewed in their distribution—with a small number of papers receiving a large share of the total citations—the ARC offers a useful snapshot of overall performance but can hide important underlying nuance. For this reason, Science-Metrix proposes to complement the ARC with the highly cited papers measure, see below.

Highly cited papers

Highly cited papers (HCP) are publications that have received RC scores among the highest in their respective field. This indicator is frequently used to examine research excellence, measuring how many high-impact articles are produced by a given research entity, relative to their expected contribution to world-leading research. For the present study, contributions to the top 10% of publications will be measured.

The HCP measure is normalised to 1, meaning that an entity with an HCP over 1 contributes more than its expected number of highly cited papers, an entity with an HCP below 1 contributes fewer than its expected number of highly cited papers, and an entity with an HCP near 1 contributes close to its expected number of highly cited papers.

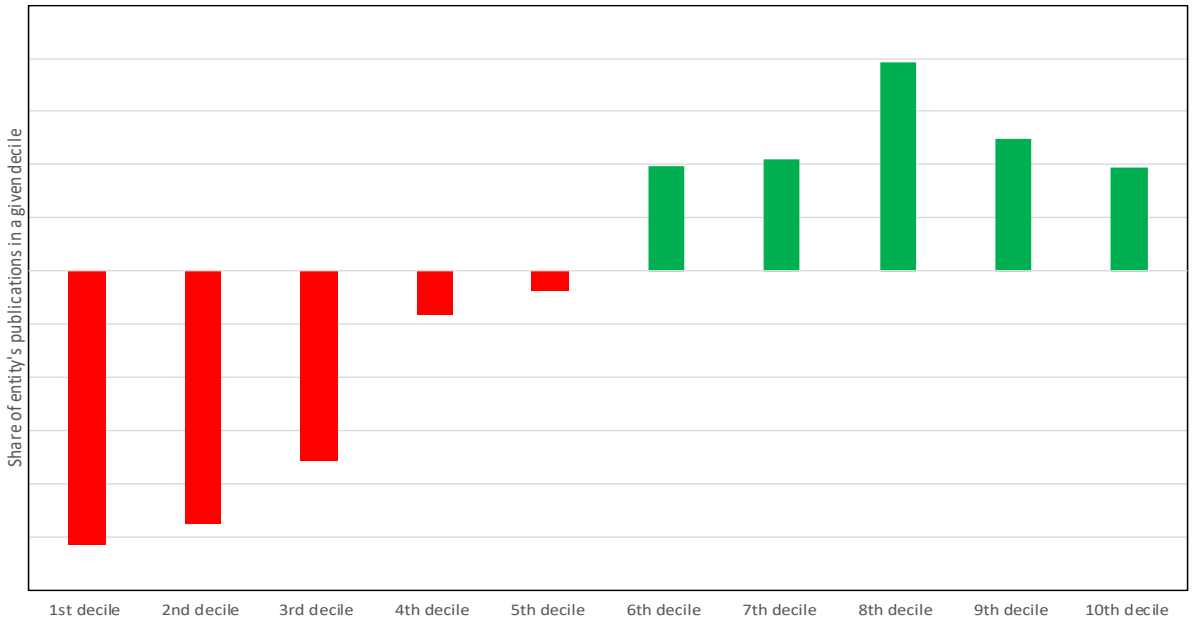
Citation distribution charts & citation distribution index

The citation distribution chart (CDC) is a tool that facilitates a simple but nuanced visual inspection of an entity's research impact relative to worldwide performance. To prepare these charts, Science-Metrix divides all publications in a given research area into 10 groups of equal size, or deciles, based on their RC scores. The 1st decile contains the 10% of publications with the lowest RC scores; the 10th decile contains the 10% of publications with the highest RC scores.

For a given research entity, it is expected that the RC scores of its publications will follow the global distribution, with an equal number of publications falling in each of the deciles. The CDC for a given entity compares that entity's scientific impact to the global level by showing how its performance compares to the world level in each of the deciles.

As shown in Figure 1, the CDC shows 10 colour-coded bars for a hypothetical entity; each bar represents the relative presence of this entity's papers in each corresponding decile. The world level, in contrast, is represented by the central horizontal line, with no bars, as it represents the uniform distribution of all the publications across the 10 deciles. Thus, the bar's colour shows whether the specific entity has more or fewer publications in that decile than expected (i.e., the horizontal line). Green bars denote production exceeding expectation in that decile, red denotes production below expectation in that decile, and the length of the bar shows how far above/below expectation the entity is in that decile. Consequently, the longer the red bar, the fewer publications are found in that decile relative to expectation. Conversely, the longer the green bar, the more publications are found in that decile, again relative to expectation. When a decile has no bar associated with it, the entity's performance is exactly in line with the expectation based on global performance. Accordingly, a CDC with no visible bars shows that the entity in question has 10% of its papers in the 1st global decile, 10% of its papers in the 2nd global decile, and so on, which, as previously noted, corresponds to the world distribution of papers based on their RC scores.

Figure 2 Sample citation distribution chart



Source: Prepared by Science-Metrix

Ideally, one would hope to be over-represented in the highest deciles, where the most impactful publications are found; similarly, one would hope to be under-represented in the lowest deciles, where the least impactful publications are found. Thus, strong research performance is shown by long red bars on the left of the CDC and long green bars on the right of the graph. In contrast, weaker research performance is depicted by long green bars on the left side (indicating more publications than expected in the less impactful deciles) and long red bars on the right side (indicating fewer publications than expected in the more impactful deciles). Figure 2 presents distributions related to best-case and worst-case scenarios.

Figure 3 Various scenarios of citation distribution charts and their citation distribution index

	CDC	CDI
Best-case scenario		50
Typical good-case scenario		25
Typical bad-case scenario		-25
Worst-case scenario		-50

Source: Prepared by Science-Metrix

The content of the CDC can also be summarised numerically using the citation distribution index (CDI). For each decile, the performance of a given research organisation is compared to the global average, and this ratio is then multiplied by the weight corresponding to that decile (negative weight for deciles 1 through 5, positive for 6 through 10). Once a score has been produced in this fashion for each decile, they are summed to calculate the CDI for the research organisation. Thus, having a higher-than-expected number of publications in the 1st decile (i.e., the lowest-impact decile) will reduce the CDI more than having a higher-than-expected number of publications in the 2nd decile. The CDI ranges from -50 (worst-case scenario) to 50 (best-case scenario), with 0 representing parity with the world level. Compared to mean-based normalised citation metrics, the combined use of the CDC and CDI makes it possible to provide reliable citation metrics even when dealing with entities that have produced few publications (from 10 to a couple of hundred).

Average of relative CiteScores

The CiteScore used by Science-Metrix is calculated at the journal level as the total number of times peer-reviewed papers published in the journal in years X-1 and X-2 were cited in year X, divided by the total number of peer-reviewed papers appearing in the journal in years X-1 and X-2. As a result, using the CiteScore to evaluate individual research publications (or the entities producing them) is equating the quality of research with the quality of the journal in which they are published.

In brief, the CiteScores of papers are calculated by ascribing to them the CiteScore of the journal in which they are published, for the year in which they are published. Subsequently, to account for different citation patterns across fields and subfields of science, each paper's CiteScore is divided by the average CiteScore of the papers published in the same year in its subfield to obtain a relative CiteScore. The final indicator computation of a given entity is the average of its relative CiteScores.

Collaboration

International collaboration rate

An international co-publication is defined as a publication that was co-authored by individuals from at least two countries. The international collaboration rate (ICR) of an entity is simply a measure of how many of its articles are co-published with international partners as a proportion of the given entity's total output. The ICR is obtained by dividing the number of international co-publications of an entity by its total number of publications (both national and international).

Inter-sectoral co-publication rate

This indicator shows the proportion of an entity's papers that are published in collaboration (i.e., co-published) between economic sectors. For this study, Science-Metrix will consider the academic, governmental, private, and NGO sectors as potential categories for inclusion in the analysis. The final selection will be dependent on the availability of a large enough volume of data. The inter-sectoral co-publication rate is obtained by dividing the number of inter-sectoral co-publications of an entity by its total number of publications.

Network analysis

Scientific research is a communal undertaking: collaboration is an important conduit for introducing new perspectives into research, highlighting assumptions, outlining new hypotheses, and sharing testing and analysis methods. The dynamics of this interconnected research ecosystem can be analysed at the level of individual researchers and institutions, as well as at the network level, where one can find emergent properties that explain features of the ecosystem.

Network-level analyses require a delineation of actors (nodes) and types of connections between them (edges). With nodes and edges defined, numerous social network analysis tools can be applied to discover the dynamics of the system. Science-Metrix has a wealth of experience working with these tools. Below is a summary of possible network indicators, some of which might be used in this project; preliminary analyses are recommended to demonstrate the potential value of the following:

- Degree, betweenness centrality, and eigenvector centrality. These indicators characterise the network properties of individual nodes therein.
- Average degree, heterophily, and homophily. These indicators characterise properties of the network as a whole.

Shares of publications with a first-time collaboration

Science-Metrix's implementation of the Scopus database enables the retrieval of the publication history of the researchers of interest as part of an investigation or evaluation. From the list of publications so obtained, it is possible to establish a list of co-authors on each researcher's prior and current publications. By identifying publications that were written with a given co-author for the first time, it is possible to identify novel scientific collaborations. Trends in the shares (%) of publications containing such new collaboration dyads provide a measurement of the collaboration intensity in a group of researchers or institution.

Disciplinary diversity

Interdisciplinary integration

Examining the material that is cited in a paper offers a reflection of the intellectual content that is being integrated in the underlying research. Accordingly, the integration of material drawn from across disciplinary boundaries is assessed through citation behaviours. The interdisciplinary integration indicator considers (a) the number of different subfields that are being cited, (b)

the distribution of those citations across the cited subfields, and (c) the intellectual proximity of those subfields to one another.

For example, a paper that draws on knowledge from four different subfields would have a higher interdisciplinarity score than a paper that draws on only three. Similarly, a paper that cites one subfield 90% of the time and the other subfields only 10% of the time would have a lower score than a paper that cites its various subfields in roughly equal measure. Finally, a paper that integrates knowledge from biology and from chemistry would have a lower score than a paper that integrates knowledge from biology and the performing arts, because the former pair is more intellectually proximate than the latter pair.

Highly interdisciplinary papers

For this study, the indicator to be computed shows what share of an entity's papers fall within the top 10% of highly interdisciplinary papers in the world (HIP; structurally similar to the HCP, see above), with each paper's interdisciplinarity score adjusted to the average of all papers worldwide published in the same subfield and same year (similar to the RC, see above).

Multidisciplinary integration

For this study, the index of Multidisciplinary integration (MI) relies on Science-Metrix journal-based classification of science. It reflects the diversity of prior disciplinary background of a paper's co-authors. It is computed by adapting the metrics by Porter & Rafols to the disciplinary profile of co-authors in a paper.[61] MI was designed to increase for teams involving authors from different subfields, particularly where these subfields are not frequently connected in Scopus. It is normalized by the paper's subfield and year to avoid coverage biases.[62]

A paper co-authored by authors whose previous papers were distributed across subfields of science in a similar pattern (i.e., having similar relative frequency across subfields), would score lower than papers bringing together authors with different background (as measured by the subfields from their prior publications), even if each of those authors, individually, have published in a less diverse set of subfields. In other words, it is the differences between the background of each co-author that increases MI and not individual authors with diverse backgrounds. Nevertheless, authors having diverse backgrounds may be more likely to increase the MI of one paper, but only if this diversity is sufficiently different from the subfields of the remaining authors. As a result of this approach, a single-author publication, no matter the diversity of is author's background, will always receive the minimum score, since the indicator is intended to capture diversity across different authors.

Highly multidisciplinary papers

In this study highly multidisciplinary papers ($HMP_{10\%}$) is employed as a complementary indicator to (MI). It is based on the (MI), reflecting the share of papers for a given entity that lies among the 10% most multidisciplinary papers in the respective subfield and year. It allows reducing the potential effect of outliers in (MI).

---

[61] Porter, A., & Rafols, I. (2012). Interdisciplinarity: Its Bibliometric Evaluation and Its Influence in Research Outputs, 20, p. 21. Retrieved from http://sites.nationalacademies.org/DBASSE/cs/groups/dbassesite/documents/webpage/dbasse_072694.pdf.

[62] Campbell, D. et al. (2015). Application of an "interdisciplinarity" metric at the paper level and its use in a comparative analysis of the most publishing ERA and non-ERA universities. *20th International Conference on Science and Technology Indicators*. Retrieved from http://science-metrix.com/sites/default/files/science-metrix/publications/campbell_et_al_sti2015_short_paper_final_web.pdf.

Altmetrics and webometrics

Although a distinction is made below between altmetrics and webometrics, please note that they are collectively referred to as altmetrics elsewhere in this report.

Altmetrics is composed of a set of methods to measure the visibility of peer-reviewed scientific publications on social media, looking mainly at mentions of these items on Twitter, Facebook, and Mendeley, and in blogs. These mentions are usually tracked through document identifiers such as DOI, PMID, and the URL of the article. The bibliometric scientific community has invested much attention and effort in the development of these indicators in the hope that they may provide improved measurements of the societal impact of science. The value of these mentions, given how general a "mention" is as a category, is hard to interpret meaningfully on its own. It might be argued that the audience for the discussion of scientific findings on social media is made up of scientists, rather than or as much as of the lay public. Additionally, it should be kept in mind that members of a research team may themselves refer to their own research on their social media pages. In this case, altmetric "citations" are more representative of self-promotion activities than broad societal uptake. Disaggregating altmetric citations by source and using different metrics within each source may help to distinguish between cases of self-promotion and uptake. Mentions in news outlets or on Wikipedia can more safely be assumed to amount to broad uptake, for instance. Similarly, indicators based on highly mentioned publications may be less sensitive to self-promotion, compare with indicators of share of publications mentioned on a given source.

Webometrics (also referred to as custom altmetrics here), instead, refers to work made to retrieve web citations (mentions of URLs, and sometimes, hyperlinks) made towards a website of interest (project websites, but also the website of a YouTube video, to take just some examples). Websites of interest could include dissemination and public engagement platforms, and URL citations would serve as a proxy measure of uptake of such tools and associated knowledge within a broad public. Web citations are retrieved either by using Google or Bing queries, specialised tools that access Google or Bing APIs, or by parsing the results of web crawlers (the latter method generally not being feasible in an ethical manner, however, as most websites prohibit parsing by crawlers). Web citations tend to have good recall but lower precision; because the URLs of science-related websites may be captured and posted on nefarious crawling websites, reference may be made to the prior content of a URL that predate the use that is of interest to the analysis. Additionally, large portions of web citations to project websites, for instance, are made by the teams associated with these websites as part of their legitimate promotional activities. Therefore, it is uncertain to which extent such web citations capture uptake as opposed to constituting internal promotion. Webometrics of non-publication outputs are therefore best considered as measuring "web presence" rather than uptake, with uptake amounting to a discrete component of the more encompassing dimension.

As a relatively new research area, prudence is advised in the analysis and conclusions based altmetrics indicators. Triangulation and specialists' opinions are especially relevant in this dimension of this study.

Highly altmetric-mentioned publications

For this study, highly altmetric-mentioned publications (HAP) are those among the top 10% (or 1%) most mentioned in selected media (News, Wikipedia, Twitter, and Facebook). To account for differences in the dataset coverage and in the intensity of internet use across years, this indicator was normalized by each year of analysis.

The HAP measure is normalised to 1, meaning that an entity with an HAP over 1 contributes more than its expected number of highly-altmetric-mentioned papers, an entity with an HCP below 1 contributes fewer than its expected number, and an entity with an HAP near 1 contributes close to its expected number.

Altmetric mentions index (AMI)

This indicator is based on the share of papers mentioned in selected media (News, Wikipedia, Twitter, and Facebook). Differently from highly altmetric-mentioned publications (HAP), this indicator considers equally all mentioned papers in a given year. Therefore, it is more sensible to self-promotion compared with HAP, while it may not capture the presence of highly mentioned papers. It offers a complementary perspective to HAP. If both indicators portrait similarly one entity, it should be a confirmatory sign of its influence (or lack of) in altmetric sources. Better pictures of AMI (with a modest HAP) may indicate self-promotion and high proportion of papers with low numbers of mentions among the entity's papers. The opposite situation may suggest that some papers of a given entity are very influential in altmetrics sources, in spite of a (possible) low level of self-promotion.

Share of publications cited by policy documents

This indicator shows the proportion of an entity's papers that have been mentioned at least once in a policy document, as recorded in the Overton database.

technopolis group

www.technopolis-group.com