

Sustainability-inclusive evaluation

Guidance for Global Affairs Canada to include consideration of the environment in all evaluations

Prepared by:

Andy Rowe & E. Jane Davidson

Footprint Evaluation Initiative

March 2023



FOOTPRINT EVALUATION

FOR A SUSTAINABLE FUTURE

Contents

Who and what is this guide for?	1
How the guide is structured.....	1
What do we mean by environmental sustainability?	1
Step 1: Getting agreement to put environmental sustainability on the agenda for an evaluation	3
Step 2: Identifying implicit worldviews and mindsets regarding human and natural systems	6
Step 3: Understanding the nexus between gender and environment	10
Step 4: Identifying two or three big issues to focus on	15
Step 5: Writing sustainability-inclusive key evaluation questions.....	16
Step 6: Getting to know “the place” for each of your big issues	20
Step 7: Expanding temporal and spatial scales.....	21
Step 8: Using systems concepts and frameworks.....	23
Step 9: Gathering evidence from existing and new sources.....	27
Step 10: Making crystal clear judgements about natural system impacts	30
Step 11: Communicating findings and supporting use	35
Call to Action.....	37
Rethinking Evaluation	39

Who and what is this guide for?

This guide has been written to support the Evaluation Division of Global Affairs Canada to include consideration of the environment in every evaluation they commission, undertake, oversee, or are involved in.

Particular attention is devoted to the natural system effects of trade, development, and foreign affairs policies, initiatives, programs, and other government undertakings whose primary purpose is not environmental. By far the greatest environmental damage is being done as an inadvertent side impact of economic and other human activity; clean-up efforts that are explicitly focused on the environment are vanishingly small in comparison. Therefore, the most powerful contribution of evaluation to sustainability will come from our work evaluating efforts that primarily focus on improving human systems.

The guide is designed to be user-friendly and to provide cost-effective guidance, helping evaluation teams to identify two or three issues worth highlighting and finding reasonably quick and cost-effective ways of addressing them within the context of a wider evaluation.

How the guide is structured

This guide is presented as key steps for integrating environment into all evaluations. The general logic for the order of presentation is to begin with the more foundational and broad reaching steps proceeding to still critical but more specific steps. Where an important consideration is discussed in a subsequent step, reference is made to that step.

What do we mean by environmental sustainability?

In this guide, we use the Footprint Evaluation definition of environmental sustainability:

“The ongoing ability of natural systems to support the equitable life of all species on earth.”

Environmental sustainability includes mitigating climate change and addressing other environmental challenges, including biodiversity loss, over-exploitation, pollution (air, water, soil), deforestation, invasive plant and animal species, access to fresh water, and the restoration of natural systems.

Natural systems globally are approaching thresholds where the critical systems will be so impaired that the sustainability of human and other life is at risk. This is a critical contextual factor and gives rise to the growing consensus that globally we need to end causing harm to natural systems. This means that we must apply a *do no harm* standard in evaluation where the net impact of all interventions are assessed against this *do no harm* standard and evaluation learning aim to contribute to interventions steadily and quickly reducing the harm they cause.

This is consistent with a widening range of development agencies, including Canada:

“Do no harm”—Canada’s development initiatives will not pollute or degrade the environment or natural resources of its partner countries. Initiatives will aim for multiple benefits that improve the environment while benefitting other sectors, such as gender equality and economic growth.¹

¹ [Government of Canada Supporting environmental sustainability in developing countries](#) 2020-07-30

Step 1: Getting agreement to put environmental sustainability on the agenda for an evaluation

One of the first tasks for any evaluation during the project inception stage (GAC Program Evaluation Process Map steps A1, A2) is to engage with primary intended users and key stakeholders and interests to agree on the scope and boundaries for the evaluation. As discussed in [Step 6](#) below, these include representatives of interests in the human and natural systems that can affect success and those that are affected by the intervention, and they should be involved during inception.

Footprint Evaluation suggests that the “Involve” and “Collaborate” positions in the International Association for Public Participation Stakeholder Interaction Spectrum (see Figure 1) indicate the desired level of involvement of interests during inception and throughout the evaluation², these are the roles that are said to promote use of evaluation and research more broadly.

Figure 1. Engaging at the ‘Involve’ or ‘Collaborate’ levels on the IAP2 Spectrum of Public Participation is suggested by Footprint Evaluation as most conducive for evaluation use and influence

IAP2 Spectrum of Public Participation



IAP2’s Spectrum of Public Participation was designed to assist with the selection of the level of participation that defines the public’s role in any public participation process. The Spectrum is used internationally, and it is found in public participation plans around the world.

		INCREASING IMPACT ON THE DECISION				
		INFORM	CONSULT	INVOLVE	COLLABORATE	EMPOWER
PUBLIC PARTICIPATION GOAL		To provide the public with balanced and objective information to assist them in understanding the problem, alternatives, opportunities and/or solutions.	To obtain public feedback on analysis, alternatives and/or decisions.	To work directly with the public throughout the process to ensure that public concerns and aspirations are consistently understood and considered.	To partner with the public in each aspect of the decision including the development of alternatives and the identification of the preferred solution.	To place final decision making in the hands of the public.
	PROMISE TO THE PUBLIC	We will keep you informed.	We will keep you informed, listen to and acknowledge concerns and aspirations, and provide feedback on how public input influenced the decision.	We will work with you to ensure that your concerns and aspirations are directly reflected in the alternatives developed and provide feedback on how public input influenced the decision.	We will look to you for advice and innovation in formulating solutions and incorporate your advice and recommendations into the decisions to the maximum extent possible.	We will implement what you decide.

© IAP2 International Federation 2018. All rights reserved. 20181112_v1

² <https://organizingengagement.org/models/spectrum-of-public-participation/>

What are the most important talking points when making the case for including consideration for the environment in the evaluation of a policy or program that was not environmentally focused?

1. All human actions and existence depend on natural systems. [The wedding cake portrayal of the Sustainable Development Goals (see
2. Figure 2, below) is a useful graphic to show how natural systems are the foundation on which human wellbeing depends.]
3. Those natural systems have been seriously damaged already and continue to be harmed, even where there has been a move toward somewhat more sustainable practices.
4. We are plunging into the abyss; without restoration, climate, nature and equity goals are unachievable.
5. Restoration is a task too huge for environmental programs alone; we need to urgently stop the inadvertent damage and steer all policies and programs toward being restorative as well.
6. Evaluations of non-environmental programs and policies have largely ignored natural system effects³, which means that many of our overall evaluative conclusions weren't entirely valid.
7. To ensure that we stay relevant to the world's polycrises including environmental breakdown, GAC evaluations will now ensure that the most important insights regarding natural system impacts are included in all of our work, so that the decision makers we serve can make informed, sustainable choices right across our government's portfolio.

³ Debbie DeLancey and Andy Rowe (forthcoming): Sustainability-Ready Evaluation – A Call to Action. Canadian Journal of Program Evaluation

Figure 2. The Wedding Cake portrayal of the Sustainable Development Goals (SDGs)

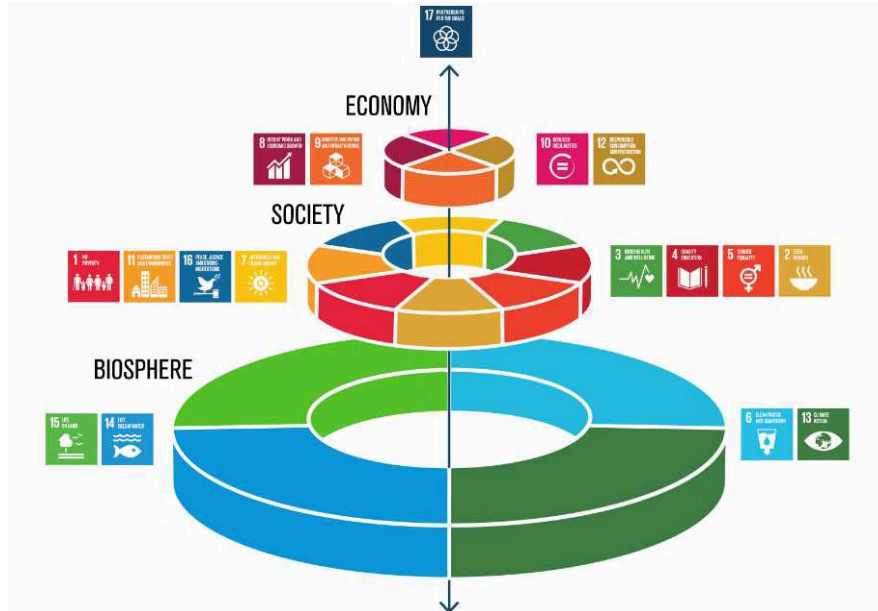
The illustration below is free to use under the Creative Commons license CC BY-ND 3.0. You are free to share — copy and redistribute the material in any medium or format.

The guidance provided to federal government drafters of Treasury Board Submissions for new, expanded or modified initiatives is very clear that environmental systems must be considered. The required preliminary Strategic Environmental Assessment (SEA)⁴ identifies six required questions about environment including the following first three:

1. The proposal has outcomes, either positive or negative, that affect natural resources.
2. The proposal has a known direct or a likely indirect outcome that is expected to have considerable positive or negative impacts on the environment.
3. The outcomes of the proposal are likely to affect the achievement of the Federal Sustainable Development Strategy's goals and targets (e.g., reduction of greenhouse gas emissions or the protection of an endangered species).

The substance of this directive can be interpreted as an important factor in the OECD DAC relevance and coherence criteria (see [Step 5](#), p. 16). However, the OECD-DAC peer learning review of mainstreaming environment in evaluation (2019, reference below) noted with concern that the mandatory requirement on environment was removed. The report made several recommendations towards mainstreaming environment in evaluation.

The following resources and sources are likely to be useful for getting the environment on the agenda



for evaluation.

⁴Section 4.2 Conducting a Preliminary Scan in [The Cabinet Directive on the Environmental Assessment of Policy, Plan and Program Proposals](#). Date modified: 2016-07-06

- [OECD \(2019\). Greening Development Co-operation: Canada Report](#). Raises concerns that mandatory requirement to cover the environment has been removed from evaluation in GoC. The report also covers: how and why environment issues (including biodiversity, climate adaptation and mitigation, and pollution) are integrated across programmes; what has worked and why; what challenges remain and are emerging; and how these challenges can best be addressed.
- [Canada's current environmental issues and international commitments](#) (World Factbook)
- Section 4.2 Conducting a Preliminary Scan in [The Cabinet Directive on the Environmental Assessment of Policy, Plan and Program Proposals](#).
- [GAC's Environment Handbook for Community Development Initiatives](#).
- GAC's Environment Framework for International Development Assistance
- [The Federal Sustainable Development Strategy](#)
- [Footprint Evaluation Webinar 2: Entry points for environmental sustainability](#)

Step 2: Identifying implicit worldviews and mindsets regarding human and natural systems

As you speak with the primary intended users of the evaluation and other stakeholders, it is important to listen between the lines for the worldviews and mindsets regarding human and natural systems that underpin the policy or program being evaluated. Depending on what you find, an important purpose for the evaluation may be to influence those worldviews and mindsets in the needed direction. In utilization-focused evaluation, this is known as conceptual use, i.e., changing the way that people frame or understand an important issue.

Worldviews

A worldview is fundamentally about what is valued, what is essential or sacred, what has status, and what is regarded as having low or no value.

The worldview that has led to the destruction of our natural environment is fundamentally an extractive, colonial view of the world. This worldview is shaped by the concept of dominion, where 'man' has ascendancy or dominion over all things. Everything else (our non-human relatives, flora and fauna, minerals, water, air, and land) has value only to the extent that it can be used to fulfil a human need or desire (economic, social, biological, etc.).

The environmentally destructive worldview came hand in hand with its socially destructive counterpart, also shaped by the concept of dominion. The so-called ruling classes (for example wealthy, land-owning white men, churches) claimed dominion over women and girls, peasants, the working classes, Black and Indigenous peoples, and people of colour. We are still struggling to dismantle the patriarchal systems that were baked in by those in power and that continue to oppress.

To effectively transform systems and practices and address the historical harm done in both human and natural systems, decision makers will need to divest themselves of this extractive, colonial worldview and replace it with a more sustainability - and social justice-savvy one. In that worldview, natural

systems are recognized as having intrinsic value and as being the foundation of all human endeavour. Rather than the colonial concept of dominion over the natural world, its starting point would incorporate concepts of stewardship of our non-human relatives and non-living things, such as those found in Indigenous worldviews.

To help decision makers transition to this way of thinking about the world, evaluation needs to systematically apply the *do no harm* standard introduced in the environmental sustainability section above, and to do so for all evaluations, including those that do not have explicit environmental goals.

Global Affairs Canada has long recognized the gendered side of dominion and has adopted a feminist evaluation framework to ensure that issues of gender and power are highlighted in all evaluations. The well-developed antennae developed using feminist evaluation need now to be expanded to highlight the equivalents of these issues in the natural system.

Evaluation makes valuable and valued contributions when it shows the way forward. That doesn't necessarily mean long lists of recommendations, especially when it comes to advancing equity and sustainability. Acceptance of recommendations is a relatively short-term form of influence. Shifting worldviews and mindsets is a more enduring form of influence, one that will also benefit *future* policymaking and decision-making.

As evaluators, then, our most important purpose here is conceptual use – to help decision makers understand and adopt the worldview and mindsets they will need to lead their organisations, their countries, and the world away from environmentally and socially harmful practices and to repair the intergenerational and ecological harm that has already been done.

Mindsets

The worldviews described above are closely associated with certain mindsets about how decisions should be made. Again, these are often subconscious, but they are pervasive. To illustrate:

... for years the widely held belief in medical circles was that women used too many health-care resources compared to men. As a result, men were viewed as the standard for seeking health care, while women were often dismissed as hysterical or “anxious” when they sought care. “We used to think women were overutilizing health care, and men were doing it correctly,” Griffith said. “What we realized was that women were doing it better, mostly for preventive care, and men were actually underutilizing health care.” (Washington Post 2023

<https://www.washingtonpost.com/wellness/2023/04/17/mens-health-longevity-gap/>)

The most pervasive mindset we see is one in which decisions are framed as trade-offs. For example, “People are struggling, and we need to help them urgently; we don't have the luxury of also being able to deal with the environment right now.” Where this mindset prevails and when trade-offs are made, we often see the implicit values of the colonial, extractive mindset emerge, with “the economy” (and especially the desires of corporate and other influential interests) typically having primacy.

Decision makers who subscribe to a more sustainability- and social justice-savvy worldview value both human and natural systems equally and will seek win:win solutions, i.e., solutions that are restorative for the environment as well as for women and other historically marginalized groups.

Ideally, we are looking to encourage decision makers to aim for better than a “do no [additional] harm” standard by seeking to repair past damage done. Restorative solutions often start with nature and take a longer-term (e.g., 7-generation) perspective on stewardship – looking back to the natural environment that our great grandparents lived in and then working to ensure that our great grandchildren will also be able to enjoy the natural environment in a thriving state.

There is a solid evidence and practice base that pursuing win:win solutions is usually within reach and feasible⁵. Collaborative processes are key to success in pursuing win:win (see the public participation spectrum in Figure 1, p. 3). In most cases, facilitation can be effectively managed by the evaluation team, but for more complex engagements, an external facilitator can be highly beneficial.

Feminist evaluation approaches are well suited to pursuit of win:win solutions, which usually entails managing strong differences in power and influence in decision making. Footprint Evaluation advocates that all interests that are able to influence the success of the decision and all interests affected by the decision should have standing in the decision processes. Ensuring that all interests are able to effectively participate in the dialogues, and that those whose power is far less are heard and their views understood, is essential to find win:win solutions that all parties can buy into.

Identifying the worldviews and mindsets implicit in policies and programs

If you ask stakeholders directly, the vast majority will, of course, say that they value both human and natural systems and situate themselves in that worldview. The same is true with mindsets. Social desirability bias means that *every* decision maker will claim to be seeking win:win solutions, even if this isn't always happening in practice. For this reason, it's usually unhelpful to ask people directly what worldviews and mindsets they consciously subscribe to, but rather to infer these things indirectly.

There is a larger point here, too. Worldviews and mindsets exist not so much in the minds of individual decision makers and other stakeholders; rather, they tend to be woven into the fabric of the policy or program itself. Every policy is formulated, and every program designed, in the context of an organizational (GAC) and government (GoC) culture that has, over time, developed norms, values, and beliefs about the nature of the world and the nature of the work, what matters, and what the default approach to evaluation, learning, and adaptive management should be. The worldviews and mindsets that are woven into policies and programs are largely reflective of this collective consciousness, the organizational culture, rather than the thinking of the individuals involved in the design.

⁵ Lawrence Susskind writes about win:win in Lawrence Susskind (2014): *Good For You: Great For Me*. Public Affairs, New York. Win:win is a general category of approaches to negotiation that seek to satisfy the most important needs of the various parties in a negotiation. The negotiation processes are the key mechanism and are outlined in the Mutual Gains Approach to Negotiation developed by the Consensus Building Institute (CBI) that was founded by Larry Susskind (see <https://www.cbi.org/article/mutual-gains-approach/>).

GAC evaluations need to help bring the worldviews and mindsets implicit in policy and programs out of the shadows and into clear view. By focusing on the worldviews and mindsets we see implicit in policies and programs, rather than in individuals, we can hopefully turn this into an exploration together rather than making decision makers feel criticized.

To understand what worldviews and mindsets are in play, we need to listen between the lines. Table 1 lists some examples of narratives and expressions to look and listen for (in documentation or when speaking with stakeholders) that can help reveal the implicit thinking that underpins the initiative.

Table 1. Narratives and expressions indicating the likely worldviews and mindsets underpinning a policy or program

<p>Extractive, colonial, ‘dominion’ worldview</p> <ul style="list-style-type: none"> • Use of the term “natural resources” (implies that the natural system is a ‘resource’ for humans). • Prioritization of “the economy” (with corporate and other powerful interests typically given primacy) ahead of the wellbeing of people and nature. • The theory of change includes little or no consideration of natural systems (even if there is some acknowledgement that the environment is valued). • Evaluation questions are not sustainability-inclusive but focus entirely on the human system. 	<p>Sustainability- & social justice-savvy worldview</p> <ul style="list-style-type: none"> • Emphasis is on human responsibility for the environment to keep it thriving (or restore it), even when there is no immediately realisable benefit for people. • Balanced prioritization across multiple human and natural system concerns, without sidelining the needs of the natural environment or of people with relatively less power and influence. • Evaluation questions are sustainability-inclusive, e.g., looking for outcomes and impacts in coupled human and natural systems.
<p>Trade-offs mindset</p> <ul style="list-style-type: none"> • “People are struggling, and we need to help them urgently; we don’t have the luxury of also being able to deal with the environment right now.” • We are addressing the value chain needs of smallholder farmers in these Districts • Evaluation questions <i>may</i> include mention of the environment, but their wording or operationalization implies that there is a certain amount of environmental damage that is acceptable in order to meet human needs, especially economic outcomes. 	<p>Win:win mindset</p> <ul style="list-style-type: none"> • The theory of change recognises points of nexus of human and natural systems and there are deliberate actions to avoid having the intervention contribute to harm in either. • The reach of the intervention is considered within its human and natural system scales. Natural system scales are, minimally, ecosystem(s) and human system scales’ reach to important social boundaries such as kinship, culture, traditional practices and so on. • Longer term solutions are advanced with shorter term approaches to ensure that harm is mitigated while the intervention is maturing.

Government of Canada worldviews and mindsets – progress to date and moving forward

Through government-wide policies and guidance and global environmental commitments, the GoC has ascribed value to natural systems, importantly influenced by the overwhelming evidence of the sustainability crisis. However, mindsets have not kept pace with the shift in worldview. For example, there is a substantial gap between the intent of national policies and global commitments addressing important climate and sustainability crises and the limited progress and pace of change. This reflects the worldview – mindset connection where policies and commitments are now recognising the value of natural systems in addressing the sustainability crisis, but the tools and mechanisms used by implementing structures (such as regulations or funding mechanisms, selection of implementing partners) come from a period and context of either/or mindsets (looking for trade-offs).

This is one of the central roles of evaluation – describing, understanding and assessing gaps between intent (in terms of desired results) and what is actually being achieved. For sustainability, this translates into the policy goal and actual achievements gap, which is reflective of the worldview – mindset gap.

As with all evaluation, employing processes that are known to promote use and influence is important. Steps [9](#), [10](#) and [11](#) pick up how sustainability-ready evaluation approaches are designed to shift mindsets to align with the sustainability and climate policies and environmental commitments.

Step 3: Understanding the nexus between gender and environment

When scoping the evaluation, one important task will be identifying the most important ways in which natural systems might be affected by an intervention. This is crucial for finalizing the list of outcomes and impacts to be included in the evaluation, for updating the theory of change to include all important outcomes and nexus effects, and to feed into the evaluation questions (we will get to this in [Step 5](#)).

It is best to do this part of the scoping in two phases – generating an initial list of possible or likely natural system effects, then narrowing it down to two or three big issues (we will cover this in [Step 4](#)).

One of the most fruitful places for identifying important natural system effects is at the nexus between human systems and natural systems. This is where the environment is affecting human wellbeing (especially that of women and other marginalized groups) and where humans are affecting nature. That may sound a little abstract, so let's see what it looks like for a couple of GAC evaluations.

Example 1: Weapons Threat Reduction (WTR) Program evaluation.

In general, military and weapons are challenging to bring to a win:win discussion. This is understandable; however, there are aspects of military and weapon decisions that are quite amenable to the flexibility needed to identify options that satisfy military requirements as well as contributing to sustainability. The most obvious examples are procurement, construction, and landscape management.

Concrete and steel are environmentally harmful, have good alternatives and are a relatively easy way for less harmful environmental impacts. Military establishments often have significant land areas to which

access is restricted, modest efforts to manage those landscapes sustainably can ensure that buffer lands provide improved environmental services such as carbon sequestration, and habitats that contribute to water management such as improving aquifers, managing flood risk and water cleansing functions. Extensive grassed and mowed areas are also common in military establishments, more environmentally beneficial options are well known and should not impair the necessary military functions.

It is well known that solar and wind alternatives are being developed that will provide a win for military operations, reduce vulnerability and the need for fossil fuels. A military that achieves *do no harm* to environment and that engages in restoring natural systems in these aspects can achieve significant reduction in the environmental harm that is an inevitable companion for many military options.

Specifically on the WTR program, weapons have obvious effects on humans, especially when deployed, but what about the environment? Nuclear and biochemical weapons (and the materials used to make them) cause serious environmental contamination when deployed. That contamination can be hugely impactful for humans in the area, who can become ill or die from contaminated water, air, and land. Likewise, these effects are hugely impactful for natural systems. Inclusion in evaluations of WTR programs of avoided direct natural system consequences will significantly enhance the benefit side. This is similar to the [Footprint Evaluation thought experiment](#) on a community corrections program in Australia, where inclusion of the strong adverse natural system impacts of prisons would significantly enhance the net benefits of reducing the prison population and so the number of prisons.⁶

Another major activity with implications at the nexus is the disposal of weapons and contaminated materials when weapons are decommissioned. Again, this is something that must be assessed with both human and natural system considerations in mind. Disposal of weapons and materials needs to be secure with respect to the environment, as well as to human safety. It seems likely that environmental problems associated with disposal can affect humans and the environment; therefore, inclusion of both human and environmental consequences will strengthen the evaluation arguments for less harmful disposal, and the evaluation's recognition of benefits from enhanced disposal methods.

Other aspects of the program that might be explored for their nexus implications include construction activities associated with WTR – for example, labs, military bases, prisons, or storage structures. Finally, if the program engages in any substantial procurement, this is always an area where we can help decision makers understand the environmental implications of their decisions.

In [Step 4](#), we'll discuss how to pick two or three big issues to highlight from the list of ideas generated by exploring the nexus.

Example 2: Partnership for Gender Equality (PGE) evaluation

The Partnership for Gender Equality evaluation, like many of GAC's corporate evaluations, is a somewhat difficult evaluation to apply this thinking to because it isn't a program or policy in a traditional sense. Fundamentally, it is Canada's participation in and contribution to an international partnership that is somewhat intangible in nature. That partnership oversees the Equality Fund, which is one area

⁶ See [Footprint Evaluation \(April 2022\): Thought Experiments Report](#)

where it may be easier to see potential natural system implications, particularly when we search at the nexus between gender and nature.

Initial meetings with the team surfaced several ideas for the team to consider:

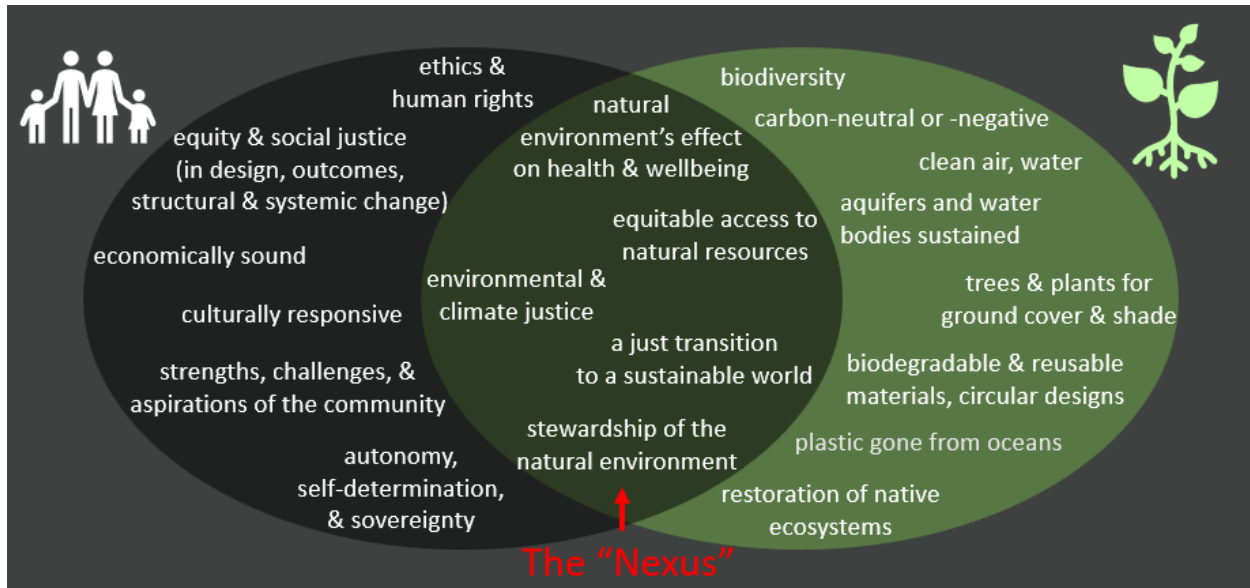
- How well is the Equality Fund directing funds toward women working at the nexus, where environmental issues are affecting or being affected by women? [Consider what proportion of the funding (a range) would be appropriate, given the importance to women of sustainability issues in the various regions covered by the Fund. How little funding would be far too little to devote to this from the total pool? 5%? 10%? 25%? How much would be overkill? Does the answer to this vary by region, population, or nexus issue? And, what is the reasoning behind these conclusions?]
- Turning up the microscope on ‘nexus’ projects (gender and environment) funded by the Equality Fund – what is the value added by projects addressing the nexus, how might they be improved, what are the lessons from where is this being done particularly well, and what promising findings are starting to emerge? Where are the opportunities to better support women addressing issues at the nexus?
- Locally led development is an emphasis in many Equality Fund projects. Therefore, a helpful framing would be to consider how well this is being done *with the natural system in mind*. In other words, how resilient to climate change are the project benefits to women – and to the natural systems that strongly affect women – here?
- The Equality Fund’s impact investing framework – does it cover environment as well as social impact, and if so, how well?

How to identify nexus issues

What tools, frameworks, and approaches might you use to identify the most important nexus issues for an evaluation you are working on?

A useful starting point is the following diagram (see Figure 3), which highlights the main ‘values’ relevant to human and natural systems, including the overlap areas at the nexus. By ‘values’, we mean the things, actions, qualities, or outcomes/impacts that tell us which policy or program features, outcomes, and impacts are beneficial or detrimental, valuable or a minus, and how well the various needs, interests, and concerns of people and nature have been balanced.

Figure 3. The main 'values' relevant to human and natural systems



Source: Davidson, E. J. (forthcoming). *Evaluation Methodology Basics: The nuts and bolts of sound evaluation* (2nd ed.). Used with the author's permission.

Using the above diagram as a conversation starter, meet with GAC's in-house environmental experts. They can help your evaluation team identify the most important nexus issues, as well as any other natural system considerations that could potentially be covered in the evaluation. They will also be invaluable in [Step 4](#), where we whittle this list down to two or three major issues to focus on. Their expertise will help guide you in deciding what's most important.

The environment team also offers short training courses to bring you up to speed on key issues, and they can point you to a wide range of tools to assess the aspects you decide to include. For many issues, secondary data sources or existing studies are available, so that the evaluation team may not need to gather primary evidence to provide a useful assessment of the issue.

Several other useful resources are available and recommended:

- [Understanding the nexus of human and natural systems](#). Footprint Evaluation brief explanation.
- [Footprint evaluation webinar 1: Identifying points of nexus between human and natural systems](#)
- provides a useful overview the "Gender and Environment Nexus" and links to a range of valuable resources
- One of the resources the Geneva Environment Network points to is this [OECD 2021 document on Gender and the Environment](#) "Gender equality and environmental goals are mutually reinforcing, with slow progress on environmental actions affecting the achievement of gender equality, and vice versa. Progress towards the Sustainable Development Goals (SDGs) requires targeted and coherent actions. However, complementarities and trade-offs between gender equality and environmental sustainability are scarcely documented within the SDG framework.

Based on the SDG framework, this report provides an overview of the gender-environment nexus...”

- [Global Gender and Environment Outlook \(GGEO\)](#). This document from UN Environment draws on the extensive UNEP network to provide a synthesis of the gender-environment nexus “by exploring future sustainability pathways from a gender perspective, we can envisage the future we can have and make a tangible difference in the lives of people around the world, while taking care of the environment.”
- Ravera, F., Iniesta-Arandia, I., Martín-López, B. *et al.* Gender perspectives in resilience, vulnerability and adaptation to global environmental change. *Ambio* **45** (Suppl 3), 235–247 (2016). <https://doi.org/10.1007/s13280-016-0842-1> “In this editorial paper, we explain the background setting, key questions and core approaches of gender and feminist research in vulnerability, resilience and adaptation to global environmental change”.
- [Gender and the environment: What are the barriers to gender equality in sustainable ecosystem management?](#) This blog from the International Union of Concerned Scientists (IUCN) addresses three topics of the gender-environment nexus: 1. Unequal and insecure rights over land, 2. Underrepresentation in natural resource decision making and leadership and 3. Gender-based violence. The latter topic is based on IUCN research and is an especially valuable contribution.
- [GAC’s Environment Handbook for Community Development Initiatives](#). Includes a great list of things to consider in the biophysical environment, along with practical advice on how to assess them, as well as mitigation strategies.
- [GoC Tip Sheet on Strategic Environmental Assessment \(SEA\): Gender Equality and the Empowerment of Women and Girls](#).
- [GEF Open Online Course on Gender and Environment](#).
- The Role of Hybrid Methodologies in Understanding Complex Environmental Issues and Promoting Social Justice. Hesse-Biber, Sharlene *International Journal for Transformative Research*, v6 n1 p20-26 Dec 2019
<https://eric.ed.gov/?redir=https%3a%2f%2fdoi.org%2f10.1515%2fijtr-2019-0004>
- [Footprint evaluation case study: Evaluation of environmental sustainability aspects of a national strategy](#). The case explored possible methods and processes for addressing environmental sustainability in an evaluation at a national scale and some of the factors that support or hinder applying these methods and processes.
- [Footprint evaluation: Thought experiments](#). The thought experiments involved revisiting a past, real-life evaluation and walking through how this could have been done differently to incorporate considerations of environmental sustainability.
- [UNEG EPE 2021-2022: Integrating Environment into Evaluation](#). A 90-minute webinar featuring leaders from the UNEP Evaluation Office, GEF, IFAD, and UNIDO.

Step 4: Identifying two or three big issues to focus on

In [Step 3](#), we explored the most fruitful avenue for identifying potential natural system effects – by looking at the nexus between human and natural systems, which is where the environment affects human well-being and where humans are affecting the environment. Think of that as a focused brainstorming exercise to unearth the possible natural system aspects to focus on in an evaluation.

Now we need to identify the two or three most important things from that list. These will be what we use to take the natural system into consideration for this particular evaluation. It is important to not expect an intervention to solve all of the human and natural system problems, nor for an evaluation to cover every possible aspect or to always promote ultimate or best solutions.

Why only two or three issues?

This guidance has been written to support inclusion of the environment in the evaluation of programs, policies, and other efforts that were primarily aimed at improving life for humans – economically, socially, culturally, educationally, etc. Covering many more of the identified issues – a dozen, for example – is inadvisable for two main reasons.

First, as mentioned above, the bulk of decision-makers' attention will be focused on how well the most important *human system* needs and aspirations are starting to emerge. We don't want those using the evaluation to feel that its main purposes have been crowded out by sustainability concerns, even quite important ones. That would likely cause irritation, confusion, or both. Instead, we will need to strike a balance.

Second, our primary concern with the natural system elements is conceptual use – influencing the way that decision makers understand or frame the relevant issues. Highlighting a dozen natural system issues (on top of multiple findings pertaining to the human system) makes it less likely that the issues highlighted will stick in their memories. If we highlight two or three, we have a much better chance.

How do we choose which issues to focus on?

To help select which two or three natural system issues to focus on in an evaluation, it will be important at this stage to get some input from someone with expertise in environmental science or a related field such as sustainability science. Global Affairs Canada is particularly lucky (and wise) to have an environment specialist team in house, which can offer expert advice as well as a range of tools that can be used to evaluate natural system effects.

As you engage in this discussion with your experts, consider the following:

- Issues with the most important environmental implications or impacts (actual or potential); these could be important because they affect critical or threatened parts of the ecosystem and/or because they are large in magnitude.
- Issues at the nexus, particularly where gender and equity are a major factor (e.g., natural system damage disproportionately affecting women and other marginalized groups; interventions addressing both equity and sustainability that were passed over in favour of a more strongly equity-focused approach).

- Issues that are relevant not just for this policy or initiative but have far-reaching implications for thinking about equitable and sustainable policy and government action in various domains.
- Issues that are currently high on the radar within GoC, other countries or internationally.

The two or three issues selected don't necessarily need to be wholly within decision makers' control to change. Sometimes it is important to raise an issue in order to broaden decision-makers' thinking – this is what we mean by conceptual use. Even if it's not possible to apply that thinking immediately to make improvements to the *current* policy or program, we are aiming for decision-makers to take these new understandings into *future* policymaking and programming. This applies well to the WTR evaluation, for example. In that way, the influence of the evaluation is intended to be more long-lasting and far-reaching than a management response to recommendations regarding the current policy or program.

What do we do with these issues once we've found them?

Once you have identified the most important aspects of the natural system to include in the evaluation, do the following:

- Revise the theory of change to include coupled human and natural systems, highlighting where, why, and how (i.e., through what causal mechanisms) natural system effects are likely to emerge, as well as how they are likely to affect human wellbeing, especially for women and other marginalized groups. [See A2 in the GAC Program Evaluation Process Map.]
- Write a set of sustainability-inclusive evaluation questions that incorporate consideration of those aspects, alongside effects on people and human systems. [We cover how to do this in [Step 5](#), below.]

As we noted, our ordering of steps might not align with the often more fluid processes in the inception phase. Thus, for example, identification of the most important aspects might occur during theory of change discussions or while formulating sustainability-inclusive evaluation questions. Sometimes, too, with collaborative approaches such as feminist evaluation, important matters can emerge later, such as during information gathering, identifying highly beneficial (modest) enhancements to the approach on the fly.

Step 5: Writing sustainability-inclusive key evaluation questions

Having identified the main natural system issues to be included in the evaluation, the next step is to write the key evaluation questions (KEQs) that will guide the entire evaluation. Typically, an evaluation will have a small number of high-level KEQs (typically 4 to 7 questions), each broken out into a series of sub-questions. These questions, especially the high-level ones, will be included in the inception report (Step A2 of GAC's evaluation process map) and then fleshed out in more detail in the evaluation matrix (Step A3).

The high-level key evaluation questions (KEQs), as well as the vast majority of the sub-questions, should be *explicitly evaluative*. This means that they should ask not just *what* the results are (a descriptive, non-evaluative question) but *how good, bad, strong, weak, beneficial, or detrimental* they are. A quick

example is shown in Table 2, which contrasts a purely descriptive, human systems-focused evaluation question with one that is explicitly evaluative and covers both human and natural systems.

Table 2. Good evaluation questions are explicitly evaluative (“How well ...?”) rather than purely descriptive (“To what extent ...?”); they also include wording that considers effects in both human and natural systems.

From these kinds of questions ...

To what extent does the program meet the most important needs of women and girls in [insert relevant contextual information]?

... to something more like these

How well does the program meet the most important needs, **in both human and natural systems**, in ways that enabled both to thrive – particularly women, girls, and other marginalized groups (including their multiple intersectionalities) *and* the parts of the relevant ecosystem(s) that are most critical or threatened?

Why is this important? Explicitly evaluative questions encourage explicitly evaluative answers. On the environment, we need to be crystal clear with decision makers about how harmful or beneficial those actual or potential effects are, rather than assuming that they can gauge this for themselves. This is critical for conveying a clear sense of urgency, including what needs to be addressed first and fast.

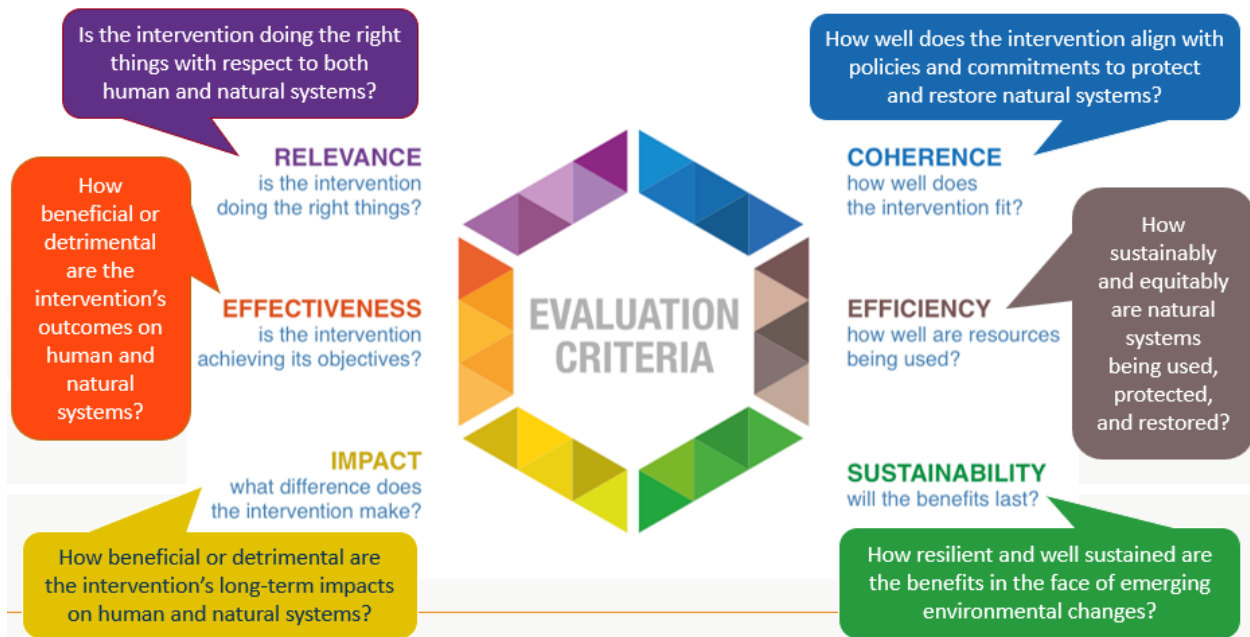
To help with the formulation of these kinds of questions, two resources are highlighted here:

- Footprint Evaluation guide to using the OECD DAC criteria to address environmental sustainability
- A generic set of Key Evaluation Questions (KEQs) to guide Footprint Evaluations

Resource #1: Using the OECD DAC criteria to address environmental sustainability

First, for evaluations that would typically be framed using the OECD DAC criteria or similar, the Footprint Evaluation initiative has developed a brief guide to framing those questions/criteria in a sustainability-inclusive way. Figure 4 provides an overview, which includes reworded versions of the short-form questions associated with each OECD DAC criterion. More details about what to consider under each one are outlined in a Footprint Evaluation resource entitled: [Addressing environmental sustainability through the OECD DAC criteria for evaluation of development assistance](#).

Figure 4. A sustainability-inclusive reading of the OECD DAC criteria's short-form evaluation questions



Note how the short-form questions in Figure 4 (in coloured call-outs) are explicitly evaluative, unlike the originals, which were mostly descriptive. For example, for the Impact criterion, instead of asking “*What difference does the intervention make?*” (a purely descriptive, non-evaluative question), we ask, “*How beneficial or detrimental are the intervention’s long-term impacts on human and natural systems?*” In other words, we are asking not just *what* happened but *how good or bad* that was – and we specify that we are considering this question with respect to both people *and* the environment.

Ensuring that your KEQs are explicitly evaluative is essential to ensure that the evaluation delivers clear answers about the implications for natural systems. If answers are purely descriptive, without clear conclusions about how good or bad the results are, we will fail to convey the needed sense of urgency. More detailed examples of good explicitly evaluative questions are available in the next resource.

Resource #2: Key Evaluation Questions (KEQs) to guide Footprint Evaluations

The OECD DAC-related resource is a brief document that deals with evaluation questions at quite a high level. For more detailed guidance, we offer an additional resource, also from the Footprint Evaluation initiative. This will be particularly useful for generating useful sub-questions as part of the evaluation matrix because it unpacks some of the details underlying each question.

The resource is a set of generic KEQs that can be adapted for any evaluation – simply reword as appropriate (e.g., replace “the evaluand” with the name of the policy or program being evaluated; specify the populations/sectors that the policy or program serves; reword using language that makes sense to stakeholders).

This resource may be used in tandem with the OECD DAC criteria, or it can be used by itself.

The seven high-level generic KEQs in the resource are shown in Table 3. Each of these questions is unpacked and explained further in the full 6-page resource, entitled: [Key Evaluation Questions \(KEQs\) to guide Footprint Evaluations](#).

Table 3. A generic set of sustainability-inclusive KEQs that can be adapted for any evaluation

1. Relevance & coherence	How relevant is the evaluand to the issues facing the population/sector and the natural environment – and how well does it complement other related efforts in the context?
2. Design & adaptation	How well does the design address the strengths, needs, and aspirations of both human and natural systems – in ways that are equitable, restorative, and enable both to thrive?
3. Implementation	How well has the evaluand been implemented so that the right people and natural system elements receive what is most needed at the right times and places and in the right ways?
4. Outcomes & impacts	How good, valuable, and important are the outcomes and impacts on both human and natural systems, particularly where equity and/or previous harm needed to be addressed?
5. Patterns, outliers & links	How did the evaluand influence change – and then how did that change continue to unfold – in the relevant coupled human and natural systems? Where, when, for whom, and under what conditions did we see the most and least valuable outcomes? Why?
6. Durability	How resilient and durable are the changes that the evaluand has contributed to, and how well are they likely to last in the face of emerging environmental and other changes?
7. Overall value	How good, valuable, or worthwhile is the evaluand overall, given its relevance and coherence, design and implementation, the value of its outcomes and impacts, their durability, and what it cost to achieve them?

Please refer to the following resources for more information:

- [Addressing environmental sustainability through the OECD DAC criteria for evaluation of development assistance](#). This resource, from the Footprint Evaluation Initiative, discusses how the six evaluation criteria of the OECD DAC (Organization for Economic Co-Operation and Development – Development Assistance Committee) can be used to get environmental sustainability on the agenda for evaluations and monitoring. (9-page downloadable PDF)
- [Key Evaluation Questions \(KEQs\) to guide Footprint Evaluations](#). This set of generic key evaluation questions (KEQs), developed by Jane Davidson and Andy Rowe, may be adapted for any evaluation. They are designed to support the inclusion of environmental sustainability by embedding consideration of the environment in each evaluation question rather than adding environmental considerations as a standalone question.
- [Footprint evaluation webinar 2: Entry points for environmental sustainability](#). A 20-minute webinar recording in which Jane Davidson and Patricia Rogers discuss several ways to get sustainability on the evaluation agenda, even for projects that have no explicit environmental objectives and where there is no mention of environmental considerations in the Terms of Reference. The approaches explored include using evaluation criteria, such as the OECD DAC criteria and using ‘Footprint-savvy’ Key Evaluation Questions (KEQs).

Step 6: Getting to know “the place” for each of your big issues

Most interventions occur at a place; projects are often at one or more places such as a communities or districts; strategies find their application at places, as do policies, often through programs. A place is where the action is, where projects, plans, strategies, policies, or other types of interventions are put in motion. And at those places human and natural systems will always couple, always because human systems draw from, deposit to, and rely upon natural systems; and because natural systems everywhere are affected by human actions. So, place is where the action of development or other interventions occurs, and it always has nexus where human and natural systems couple.

Places are not just points on a map, they are locales involving physical and social attributes. Social attributes such as culture, economy, tradition, organisations, population density, inequity, race, gender, legal and governance structures, and so on. Physical attributes such as the character, status and trends relating to topography, ecosystems, biodiversity, water, and terrestrial character, and so on. The temporal and spatial scales of human and natural systems will differ widely and need to be considered (see Step 7: Expanding temporal and spatial scales, p. 21).

All human and natural systems have “interests” and in some instances these interests have gained the status of rights. Interests are similar to stakeholders, which are those taken to represent an interest, such as women’s organisations as stakeholders in development, gender, education, economic and other interventions. The difference can be important. Stakeholders are mainly representatives of interests in the human system that have found their way to having a “stake or role” in the intervention. Interests have “an interest” in the intervention or the condition that the intervention is seeking to change because they can affect success of the intervention or are or could be affected by the intervention. Interests include those in human and natural systems at the place and, more broadly, that are connected to the intervention and targeted condition, either directly or indirectly. “Indirectly” includes what evaluation refers to as unintended or unexpected, as well as what Footprint Evaluation terms as “ignored”. Interests are found in human and natural systems; they include those not directly connected to the intervention; and they frequently are not considered in the framing of the intervention.

Interests and stakeholders are importantly place-identified; interests are more place-centric while stakeholders tend to be more intervention-centric. In this way, knowing the place means knowing the place as a nexus of human and natural systems, where the human system is attempting to change a condition, usually regarded as lying in human systems, and which will cause nexus effects in human and natural systems.

Evaluation is not long-term curiosity research, it needs to be feasible and increasingly needs to be responsive, rapid and used. So, while ‘place’ is a very wide and complex concept, knowing the place for the purposes of evaluation needs to rapidly focus on the main attributes of the places where the intervention occurs and affects, directly and indirectly.

This raises the critical matter of “main attributes of the place” for who, and what? [Steps 9](#) and [10](#) address doing this for evidence gathering and evaluative reasoning. But the main things are first identified though processes of knowing the place.

The existing feminist processes employed by GAC are intended to engage those without power or with less power in evaluation processes. The natural system typically has even less power than marginalized groups because it can't speak for itself; this is why we must explicitly represent the natural system interests in all evaluations so that it, too, has a voice.

Regardless of the processes, it is critical that the understanding is of the place and the role of the intervention at that place, not just an understanding of the intervention with the place as a backdrop or a natural resource. The usual intervention-centric evaluation processes systematically demote to unexpected / unintended or just plain ignore important effects, especially those experienced by those with less power and authority and systematically including interests representing natural systems.

Processes for knowing the place should, to the extent feasible, engage interests to help the evaluation understand the many dimensions of the intervention, especially *reach* within and across systems⁷, and the importance of the multiple effects to these interests. This is a precursor to assessing the likely strength of effects on the two considerations of central importance at this time, environmental sustainability and equity (more on this in [Step 10](#), especially Figure 9, p. 33). It also helps in identifying those interests associated with sustainability and equity and so might be considered as “core” that should continue to be engaged in the evaluation, and whose views of the intervention and especially its reach and effects need to be considered as main or central matters.

Step 7: Expanding temporal and spatial scales

Natural systems operate on a range of very different temporal and spatial scales and only in the most unusual circumstance do they align with the frames used in human systems. Natural system spatial and temporal scales can be broader or narrower, longer or shorter than those that apply to the human systems to which they are coupled. For example, management of aquatic or migratory or resident species are all shaped into human management boundaries that bear no resemblance to the ranges of the species. Forested areas, and human interaction with these will be managed by multiple agencies with different and often contradictory mandates and boundaries, and with significant differences in the spatial and temporal scales for managing tree harvesting, hunting, agro-forestry, recreation, parks and reserves, conservation, protection of avian and terrestrial species, resident endangered species, and so on. These are almost always managed by and for human use of the species. A sustainability-inclusive evaluation needs to identify the spatial scales relevant for the evaluation and while they must be broader than the frame of the intervention alone and include natural system frames, it cannot take on the multitude of applicable spatial frames that could present at the nexus.

⁷ See Steve Montague, Nancy L. Porteous, The case for including reach as a key element of program theory, *Evaluation and Program Planning*, Volume 36, Issue 1, 2013, Pages 177-183, ISSN 0149-7189, <https://doi.org/10.1016/j.evalprogplan.2012.03.005>. (This article focuses exclusively on human systems but the concepts are useful for evaluating at the nexus.)

Sometimes human spatial scales can camouflage the relevant spatial scale for natural systems. For example, the ^{8*} Woodpecker is listed as endangered in Treaty 8 and unceded territory in northwestern Canada⁹, and the population is quite small below sustainable levels. The * woodpecker however is not threatened further south across the Canada US border in *. The territory of the * woodpecker includes the US areas and stretches slightly north of the Treaty 8 boundary. Because Canadian agencies are required to protect endangered species such as the * woodpecker, interventions that would affect their habitat need to provide a high level of protection for this listed species, thereby removing a number of forest management options. These options would likely only affect the location of some of the * woodpeckers, not their sustainability since they can move to other areas within their habitat. The key mechanism here is the spatial boundary of the management structures for Canadian agencies applying to a species with very different boundaries.

The difference in temporal scales is also important. Evaluation tends to occur periodically but usually within a five- to seven-year timeframe. While it is customary for evaluations to consider humans within a range of time scales and sometimes addressing intergenerational issues, humans in general have a fairly finite and known lifespan that varies somewhat over time and by factors such as economic position, gender, place (war torn or peaceful, rural or urban, rich or struggling) and others. Other species potentially affected by / contributing to interventions have lifetimes that can be expressed in centuries (forests, turtles) to decades, years or days. At nexus points where human and natural systems couple there will be a very wide range of relevant temporal and spatial scales and of course it is neither possible nor relevant to embrace them all.

Government agencies usually operate within performance management structures that are typically annual. This human system structure and its temporal scale can strongly affect performance of interventions in natural systems. For example, credit is given for removal of x km of invasive species in a year. The area cleared of invasives needs to be maintained, else the invasives will immediately return in abundance to this newly improved (for them) habitat. Maintenance lies outside the temporal frame for performance measures, and in any case credit for maintaining a site would be highly unusual.

The two examples, the * woodpecker and invasive species show the need to incorporate the spatial and temporal scales relevant for the environment at nexus. This is highly relevant for evaluation, looking in the wrong place or at the wrong time will generate wrong observations, observations that are likely to have the wrong sign.

As the evaluation processes consider the more important things to consider the relevant temporal and spatial scales also need to be decided. From the outset, these should be presumed to differ from those operating in the human systems. In evaluation, this starts with developing the story of the intervention through a theory of change or other means. Opening up the story to include natural systems and nexus requires also reflecting on the spatial and temporal scales that the evaluation should adopt. This

⁸ It is either the Downy or Hairy Woodpecker – checking which

⁹ As this guidance was being finalised an agreement was reached between governments and First Nations that included this matter. <https://www.theguardian.com/world/2023/apr/17/canada-first-nations-land-claims-dispute-settlement>

“opening up” needs to be continued throughout with evaluation questions, information gathering and analysis and communications. And it needs to include the temporal and spatial frames relevant to all the important or core outcomes at the nexus ([Step 4](#)) and to the place ([Step 6](#)).

Resources

- Rowe, A. (2012). Evaluation of Natural Resource Interventions. *American Journal of Evaluation* , 384-394.

Step 8: Using systems concepts and frameworks

The use of systems approaches in evaluation has taken off in recent years, particularly as more evaluations focus on efforts to influence change in complex systems (e.g., industry transformations, social and political movements, changes to economic systems, and environmental restoration). The literature that this movement draws from is vast, spanning multiple disciplines. In this section, we highlight just a handful of systems concepts that have proved particularly useful when evaluating emergent change in human and natural systems.

System boundaries

Fundamental to the use of systems approaches is defining “the system” for the purposes of the evaluation. We do this by defining its boundaries. Boundaries are what determine what’s “in” and what’s “out” of the evaluation scope, what counts and what doesn’t. The systems field brings particular strengths in that it explores the implications of these choices very carefully, mindful of the implications and consequences at every step.

We’ve already made the first big decision here – that our evaluation scope boundaries (e.g., the outcomes and impacts we look for) should be positioned to include natural systems as well as human systems. As we’ve already discovered in the discussion of temporal and spatial scales ([Step 7](#)), the spatial and temporal boundaries for the initiative and its locale (“[the place](#)” – [step 6](#)) need to make sense from a natural systems perspective. We shouldn’t automatically adopt human system boundaries (such as international borders, provinces, cities, or other administrative regions) or timescales (such as human lifespans or the fiscal year), which are usually meaningless from a natural systems perspective.

Another “boundaries” decision we have discussed is the inclusion of all relevant interests (see [Step 6](#)). Unlike most evaluations, we have not restricted ourselves to considering only stakeholders (humans or entities with some kind of “role” or “stake” in the initiative). Rather, we have expanded our boundaries to cover all “interests” – all people and all living and non-living things that either (a) can influence the success of the initiative and/or (b) are affected by the initiative. That includes people who are connected to the place rather than to the intervention; it also includes the land, water, air, flora, fauna, and other parts of the natural system, which usually have no voice or representation in human systems evaluation.

One of the most important aspects of deciding what “counts” is determining the values that will be applied in the evaluation. Again, we have already delved into this question and have expanded the relevant values from the usual human system ones (ethics and human rights; equity; the strengths, challenges, and aspirations of communities; sovereignty and autonomy; etc.) to also include natural system values (e.g., biodiversity; carbon-negative or -neutral; clean water and healthy aquifers; etc.) and

issues at the nexus (e.g., equitable access to natural resources; climate and environmental justice; stewardship of the natural environment) – see Figure 3 (p. 13).

There are several other topics that systems experts advocate discussing with respect to boundaries, including whose and what kinds of knowledge and expertise are honoured and which are marginalized (see also our earlier discussions about multiple knowledges and highly inclusive processes), as well as what worldviews underpin “how a system handles interests (people or things) that are negatively affected, marginalised or victimised by it”¹⁰ and the system’s moral authority or legitimacy in doing so.

Clearly, many of these ways of thinking about systems are as powerful for exploring gender and equity issues as they are for considering the natural system. A terrific resource for exploring these further is the Williams and van ’t Hof workbook,¹⁰ which is listed as a recommended resource for this section.

Coupled systems

Although we have now drawn boundaries and defined “the system” for the purposes of the evaluation, the reality is that we are dealing with multiple nested and connected systems. A single intervention will involve human and natural systems connecting and interacting at one or multiple points. The causal character of these connections is why many refer to the systems as “coupling” rather than connecting. How do these coupled systems affect each other, and what are the implications for sustainability-inclusive evaluation?

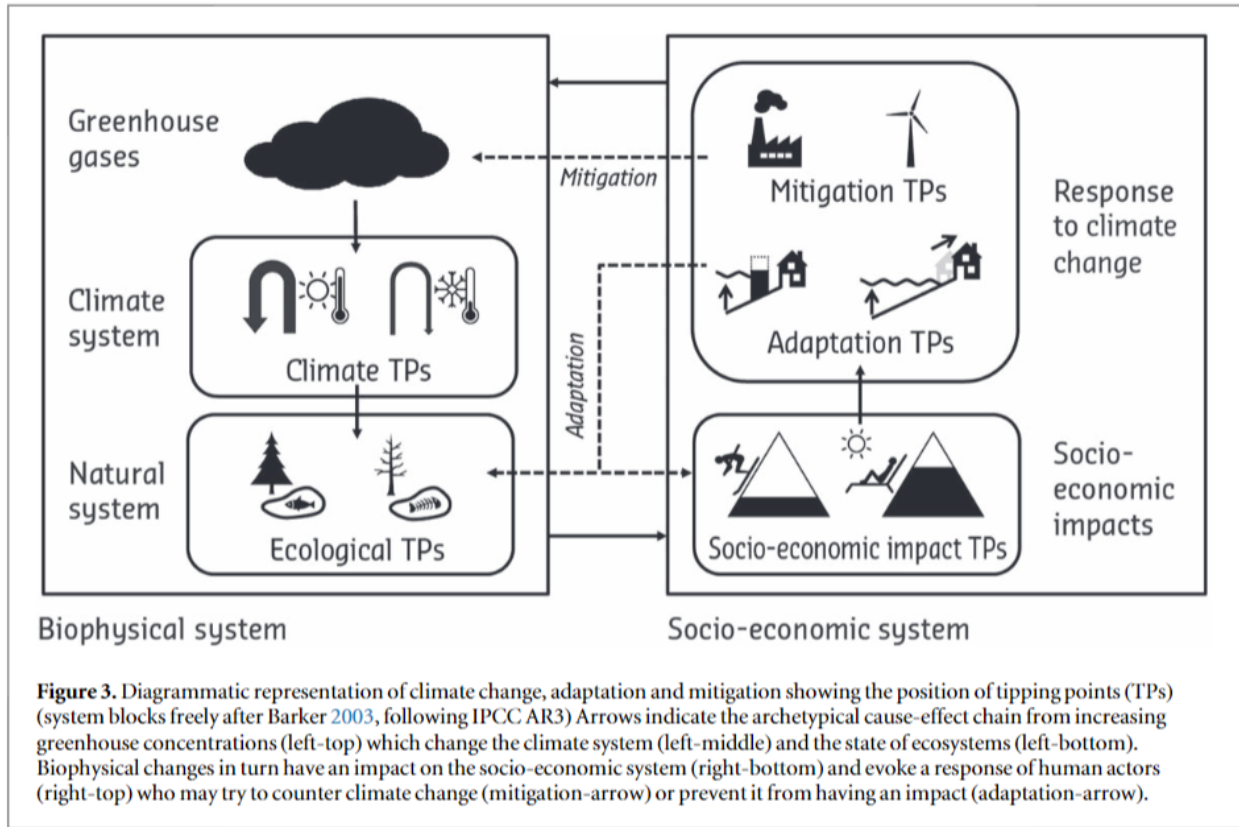
Human life is dependent on being able to draw from natural systems. This includes the necessities of air, water, and nutrition, as well as the importance of natural systems for physical, mental, and spiritual health and well-being. Beyond the value that natural systems have to humans, they also have value to other natural systems with which they couple. Damage to one system can have devastating ripple effects on other systems.

One of evaluation’s responsibilities is to recognize the burden that human actions place on natural systems and to contribute to finding and valuing alternatives.

The following diagram illustrates some of the ways in which human (socioeconomic) and natural (biophysical) systems interact and affect each other (see Figure 5). Notably missing but important to bear in mind are the *counterproductive* activities of humans that exacerbate environmental damage and slow mitigation and adaptation progress – climate change denial, disinformation, counterlobbying against the introduction of equity- and sustainability-focused legislation, relocating businesses to less regulated jurisdictions, circumventing social and environmental protections, backsliding progressive legislation or creating large loopholes, and so forth.

¹⁰ 3.3 Boundaries: critique for impact. In B. Williams & Sjon van ’t Hof (2014). [*Wicked solutions: A systems approach to complex problems*](#).

Figure 5. Natural and human systems are strongly coupled and affect each other in numerous ways



Source: *Climate change induced socio-economic tipping points (2020)*

Non-linear change

Change in complex human and natural systems does not typically emerge in a linear way, yet we often see linear expectations in policies, programs, and evaluations when it comes to the outcomes of change in such systems. For example, if a 70% reduction in a harmful outcome is needed over the next 7 years, real-time and midstream evaluations should not assume the simplistic, linear expectation that the reduction should be 10% per year. Rather than linear change, the typical pattern is an S-curve, with slow change initially until a tipping point is reached, followed by fast acceleration.

When evaluating progress, it is important to try and understand where the change is at relative to the tipping point and whether the right groundwork has been laid down to allow the needed change to pass the tipping point. In many cases, the changes we are seeking for the benefit of the environment are changes in human behaviour – especially commercial activity. Here, a useful framework is the innovation adoption curve. The tipping point for change in human systems is reached once the innovators and early adopters have started to lead the way and the early majority starts to follow the lead of opinion leaders, thereby accelerating uptake. Motivations are different for people and organizations at different points in the innovation adoption curve; this can be helpful when evaluating the influence strategies used to persuade businesses and others to change their practices.

Non-linear change is also relevant when explaining how our natural system continues to deteriorate, with multiple irreversible tipping points looming by about 2030. An important part of our evaluative work will be making it clear to decision makers how perilously close we are to these tipping points and that merely slowing down our approach is insufficient to avert disaster.

Feedback loops





All complex systems have feedback loops in the form of cause-and-effect relationships, some of which reinforce or amplify an effect while others balance or limit an effect. When evaluating outcomes and impacts in complex systems, both human and natural, we often see change emerging much faster or much slower than expected, and these patterns are often due to feedback loops. By identifying the causal loops that are slowing change or accelerating it, we can better understand and explain why and how some system effects are frustratingly slow to emerge while others seem to spiral out of control.

Feedback loops are also important for understanding whether and how we are approaching tipping points, when we might see an abrupt change as the system “flips” to a new relatively stable state. These amplifying loops are quite well understood in the natural system, but they occur in the human system as well. For example, innovators who first start inventing and using equitable and sustainable business practices are often ignored or dismissed as “fringe” until early adopters see that they are achieving success and decide to adopt these practices themselves. Some of these early adopters tend to be opinion leaders known for spotting opportunities and making well-informed decisions. Once they have joined the change, the early majority sits up and takes notice, realizes that the change is inevitable if these influential players are on board, and they also start to make change. Here’s the reinforcing loop: the more the early majority gets on board, the more the sector starts to see the change as inevitable and joins in. This reinforcing loop is what activates the tipping point and accelerates adoption.

Tipping points

As illustrated in the above example, tipping points occur when gradual change reaches a critical threshold where self-amplifying feedback loops kick in to accelerate the change, usually quite abruptly. This transforms the system into a new stable state with limited reversibility. When evaluating change in natural and human systems, we need to be able to give clarity about how far the situation is from important tipping points and how quickly or slowly we are moving toward them. Different kinds of tipping points may exhibit subtly different patterns, as illustrated in Figure 6.

Figure 6. Four kinds of tipping point and how they differ

Characteristics	Climate tipping points	Ecological tipping points	Transformation tipping points	Adaptation tipping points
				
Stable states (from state A to state B)	Qualitatively different states of a large component of the earth's climate system (A: strong ocean current; B: substantially weakened ocean current)	Distinct dynamic regimes of an ecosystem (A: oligotrophic lake state—"good" water quality; B: eutrophic lake state—"poor" water quality)	Shift in uptake of ideas, technology etc. (A: idea, technology or behaviour only for a small minority of 'early adaptors'; B: embraced by many people, on macro-level)	Formal or informal objective/performance threshold is exceeded, requiring a change in action (A: strategy 1; B: fundamentally different strategy 2)
Mechanisms	Internal system feedbacks explaining state changes as well as stabilisation	Internal system feedbacks	Diverse, mostly market mechanisms driven by human behaviour	Crossing acceptability, technical or economic threshold
Abruptness	Rapid on geological timescale and potentially relevant on a policy-relevant timescale	Rapid compared to typical change in the ecosystem	Rapid change compared to normal uptake of ideas or behaviour	Rapid change of policy or action
Other	Irreversible from human perspective Negative framing: to be avoided	Hysteresis: restoring original state through different trajectory	Positive framing: to be achieved	Can be used to construct adaptation pathways to meet objectives under changing conditions

Source: *Climate change induced socio-economic tipping points (2020)*

Resources:

- [Setting the mood for a strong global climate agreement](#) (2015 blog post explaining positive and negative tipping points in the natural system and in human behaviour change).
- [Diffusion of innovations](#) (a short book review with a succinct summary of the innovation adoption curve and how motivations to change vary across it).
- [Systems Concepts and Tools](#) (web page from Bob Williams, leading systems evaluation theorist; includes brief explanations of core concepts plus links to tools and resources).
- [Wicked solutions: A systems approach to complex problems](#) (a practical workbook from Bob Williams and Sjon van 't Hof showing how to apply key systems concepts).

Step 9: Gathering evidence from existing and new sources

[Planning for this under A3 in the flow chart (evaluation matrix).]

The possibility of systematically addressing sustainability in evaluations was remote five years ago and now real but facing challenges in going to scale. We have to assume that incorporating sustainability in evaluations will soon be an active expectation for commissioners and users, else evaluation will diminish in relevance and utility. This is somewhat analogous to evaluation systematically incorporating gender, race and Indigenous, what was unusual or ignored became accepted as necessary and then (or is becoming) routine. Along the pathway from unusual/ignored to accepted to expected the burden of proof adapted. Early on the "truth" of the importance of gender, race and Indigenous needed to be demonstrated in evaluation and more broadly. As awareness of the real truth grew, with contributions from evaluation and systematic gender, race and Indigenous inequity became accepted the focus shifted to incorporating these dimensions throughout the evaluation, questions shifted from whether something was occurring to how the interventions under evaluation stand, contributed to or reversing the conditions (see Typology in [Step 10](#), Figure 7, p. 30).

It is no longer necessary for evaluation to prove that there are very real threats to sustainability and that these result from human actions. Everyday discourse now assumes this; journalists, academics, politicians and government now take the stance that these are truths and that the matters are urgent. This strongly affects the character of the evidence that evaluation needs to present. We are currently in

a setting where the policy stance increasingly accepts the truth and urgency of sustainability and equity crises, but tends to do so separately so that policy directions on sustainability and on equity exist side by side and centrally, but generally lack concrete connections. [Step 2](#) of this guidance points to the policy / intervention gap where many interventions not aligned with policy, and points to this as an example where the worldview is shifting (policy) but mindsets are still trapped in an either-or framing. Most often this mindset is expressed as “yes we can include gender but the cost in terms of economy will limit the extent to which we can contribute to improving the condition of everyone, including women”.

The authority of this either-or mindset can be undermined with evaluative evidence that interventions are in fact harming, turning the statement into “we prioritised economy and this had the unfortunate result of harming women / Indigenous peoples / environment”. This is the stage we are currently at, where interventions usually claim to not being causing harm to natural systems but are in fact generally doing so.¹¹ To illustrate a recent thematic evaluation of support for climate adaptation by smallholder farmers from the International Fund for Agricultural Development (IFAD) found that 70% of the cases reviewed were causing net harm to natural systems, despite the organisation requiring a “do no harm” standard for projects and programs. This raises the importance of evaluation actually assessing the value of interventions in terms of harming – doing no harm – restoring, in other words the alignment of the intervention to policy (see Step 10: Making crystal clear judgements about natural system impacts).

There is direct implication for evaluation questions and evidence.

In the KEQs this guidance presents a view favouring incorporating sustainability into all questions, and not separating off into a separate criterion. The rationale is that the mindset of the intervention is likely to be misaligned to policy and guidance, hence important insights are gained through looking at the alignment and how it became misaligned as well as the actual effects of the intervention on environment and equality.

And as we have said above, the focus needs to be on conceptual use, that is in contributing to shifting the worldview and mindsets of decision makers to systematically consider and respect natural systems. Importantly because bringing natural and social sciences together can be challenging with each tending to focus on fine grained measurement when a solid approximation is fully sufficient. For conceptual use evaluation needs to address questions such as is the intervention harming natural systems – yes or no – a lot or a little. Who benefits and who suffers from this – a lot or a little? Responding to these questions requires credible and legitimate information and analysis addressing questions that are framed in ways that are salient to interests and decision makers. Co-generative knowledge processes are well suited to this.

Often, as was the case in the IFAD smallholder evaluation, solid case studies that incorporate natural system questions and which build on dialogue with interests, science knowledge holders and case study

¹¹ Project and program proposals are now required to complete an environmental assessment and to state that this project will not cause harm to environment, as do the Treasury Board requirements in Canada. Often this is a checkbox exercise and completed honestly by proponents whose expertise and experience does not support the same level of critical thinking and appraisal that they bring to human effects, and using partial or dated guidance.

authors can reach agreed assessments of the net effects of an intervention at country and local levels. Consensus seeking processes involving interests can provide quality and balanced evaluative assessments that have more authority than detailed technical judgments. At this time the key conceptual use is about addressing sustainability with urgency, not fine tuning to gain improvements with interventions that have a solid constituency.

And we should be under no illusion that evaluation resources will be immediately and sufficiently augmented to fully bring sustainability into the scope. In some instances, the evaluation resources will be supplemented, in others resources can be found within the evaluating organisation such as is the case at GAC with a strong environment division. Potentially there can be tensions with natural sciences whose methods usually involve multiple years and somewhat detailed empirical and experimental inquiry. At GAC this risk of tensions between sciences is unlikely given the strengths and approaches of the evaluation and environment divisions. We cannot anticipate how well inclusion of sustainability will be financed, but a strong evaluation culture using feminist methods partnering with a strong internal environment division suggests that resources will not be a strongly limiting factor. And the existing mixed method approaches employed by the evaluation division are fully appropriate.

The challenge is how to access the necessary knowledge of natural systems. Communities are an important source, so long as questions about natural systems are included and well framed. Additional technical knowledge about the effects of the intervention is often widely available from public sources and through targeted science knowledge holders – processes that the environment division appears well suited to support. Footprint Evaluation suggests that boundary spanners can play an important role where in-house expertise is not available. A boundary spanner will span the social and natural divide, bringing to the evaluation mature expertise from natural sciences and contributing to the design and implementation of the evaluation, for example through identification of nexus settings and of the character of these settings, suggesting appropriate technical literature and sources and assisting and facilitating the evaluators in accessing this knowledge.

At this stage there are few if any technical barriers to including sustainability and natural systems in evaluation. The barriers are political, cultural, social and economic, usually underpinned by worldviews that do not recognise the value of natural systems and enmeshed in either-or mindsets.

It is important to always be strongly aware that the evaluator is not the first to consider these matters. In our own work we have found methods such as GIS, life cycle analysis, best practices for industrial parks, ISO standards and analysis, energy efficiency standards and processes, best management practices for conservation to be invaluable. Public domain sources are valuable, especially investigative journalism; and external advisors (as consultants or from government agencies or ngos) are often helpful. Concepts such as circular or renewable economy approaches, planetary boundaries and similar ways of thinking of coupled human and natural systems can help introduce do no harm and restorative thinking about coupled systems.

Resources

International Resources Panel <https://www.resourcepanel.org/>

Rowe, A. (2019). Sustainability-Ready Evaluation: A Call to Action. *New Directions in Evaluation*, 29-48.

Rowe, A., & Lee, K. (2013, April 15). *Linking Knowledge with Action*. Retrieved August 5, 2013, from David and Luckile Packard Foundation: <http://www.packard.org/2013/04/science-subprogram-shares-new-materials-about-strategy>

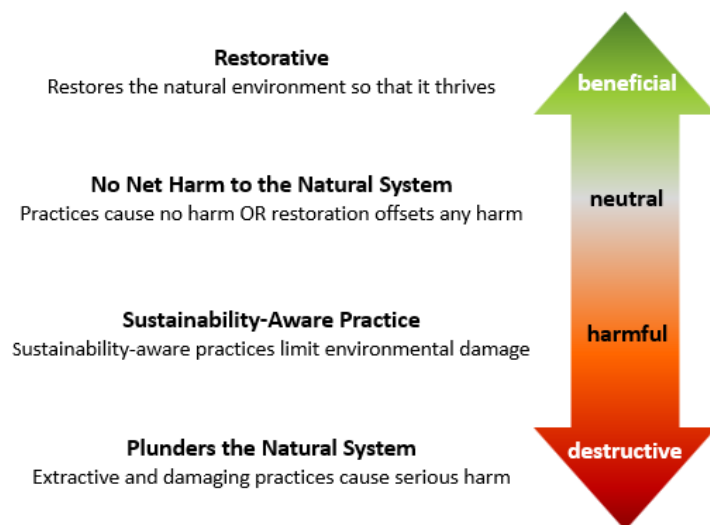
Footprint Evaluation (2023) Evaluating the environmental impact of personal protective equipment (PPE) in the COVID-19 pandemic

Step 10: Making crystal clear judgements about natural system impacts

In [Step 5](#), we mentioned the importance of writing explicitly evaluative evaluation questions, i.e., questions that ask not just *what* the results are but *how good, bad, beneficial, or harmful* they are. An important reason for this is to avoid slipping into descriptive evaluation, where results are described and the reader is left to make up their mind about how substantial and valuable any positive results are, as well as the seriousness of identified problems or harmful effects. There is a huge risk with that approach, that decision makers will fail to pick up the needed sense of urgency or the need to shift their worldview and mindsets.

To support evaluation teams in this task, the following typology was developed by Andy Rowe and has been refined in collaboration with the Footprint Evaluation team (see Figure 7). The typology helps clarify a major issue that all evaluators should be aware of. The vast majority of policies and projects are, at best, in the sustainability-aware space. This means that, although they may be less harmful than the extractive practices that have plundered natural systems for decades, they are still doing harm to the environment, just at a slowed pace.

Figure 7. A typology, or mini-rubric, for assessing actual and potential natural system effects



Source: IFAD IEO (2022) <https://webapps.ifad.org/members/ec/115/docs/EC-2021-115-W-P-3-Rev-1.pdf>; Rowe (2022) *How to address environmental sustainability in your evaluations*, National Evaluation Capacities Conference, Turin Used with permission.

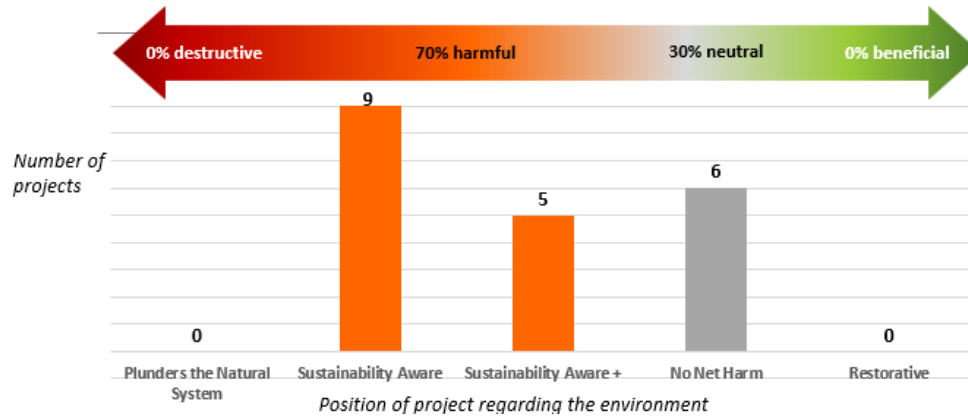
How should an evaluation team apply this typology?

The typology should be applied when answering any evaluation question relating to design, outcomes, impacts, or natural system use. For example, let's take the OECD DAC criterion of relevance and use the sustainability-inclusive version of the short-form question: *"Is the intervention doing the right things with respect to both human and natural systems?"* To answer the natural systems aspect of this question, the typology would be applied by using one of the four ratings to say clearly and succinctly whether the policy or program's design, implementation, and practices are, on balance, likely to be destructive, harmful, neutral, or beneficial with respect to the environment.

Any assessment against the typology should be supported with evidence about aspects of design, implementation, and practices that are harmful or beneficial, which are the most serious and why, and the reasoning for the overall assessment. The evaluation team would not usually need to gather primary evidence in order to draw such conclusions. The environmental sustainability of various practices are well researched already, so it is usually a matter of pointing to relevant studies and showing why and how they apply to the policy or program being evaluated. GAC's in-house environmental team will be able to point you toward the studies you will need to support your conclusions. These can include detailed quantitative studies such as Life Cycle Analyses (LCAs) right through to mixed method and qualitative studies clearly showing the destructive or beneficial effects of certain practices.

If the typology is applied to several evaluation questions, it may also be useful to synthesize these and provide an overall assessment of the project's position with respect to the natural environment. A recent study commissioned by IFAD did just that for 20 country case studies addressing climate adaptation for smallholder farmers. The study found that 70% of those were causing net harm to natural systems (although they were all "sustainability aware" in that they were making an effort to cause less harm than they might); 30% were causing no net harm (some with restorative elements); while zero were restorative overall (see Figure 8).

Figure 8. An IFAD analysis of country projects addressing climate adaptation for smallholder farmers found that 70% were causing net harm to natural systems, while none were restorative.



Source: Footprint Evaluation (forthcoming).

A clear, succinct analysis like this is far more likely to create a lightbulb moment and galvanize action. This is why we strongly advocate for providing clear evaluative conclusions, rather than leaving evaluation findings open to other interpretations.

What standard should we be applying as “good enough”?

The absolute minimum standard for all policies and programs needs to be in the No Net Harm zone. However, given the harm already done to our environment, we do need to urge decision makers to bring as many policies and programs as possible into the green zone, where past harm is repaired and the natural system restored. This is the only way we have a hope of meaningfully slowing down climate breakdown and other forms of environmental degradation.

The other way in which we need to be crystal clear about results is with respect to “status and trends.” Many natural systems are approaching irreversible thresholds.¹² It is essential that human actions harming natural system, as well as those harmful effects, are identified (status) and that the agents responsible for the interventions develop other actions that move towards and achieve win-win solutions for human and natural systems (trends).

Evaluation does not provide a useful service by saying we are getting better – that is, we are heading towards a cliff, but our speed has diminished somewhat. Incremental improvement is better than nothing, but it is critically important that we clearly convey the urgency of the situation so that decision makers are clear that congratulations are not yet in order, but rather more urgent action. Evaluation needs to say we have moved x km nearer the cliff (status) which is now y km away, and while our pace has slowed (trend), the cliff is closer and we are projected to fall off it by 20XX, therefore we need to change direction immediately (win:win). Too often, evaluation focuses on trend alone without reference to status and the consequence of the [too-slow] pace of improvement.

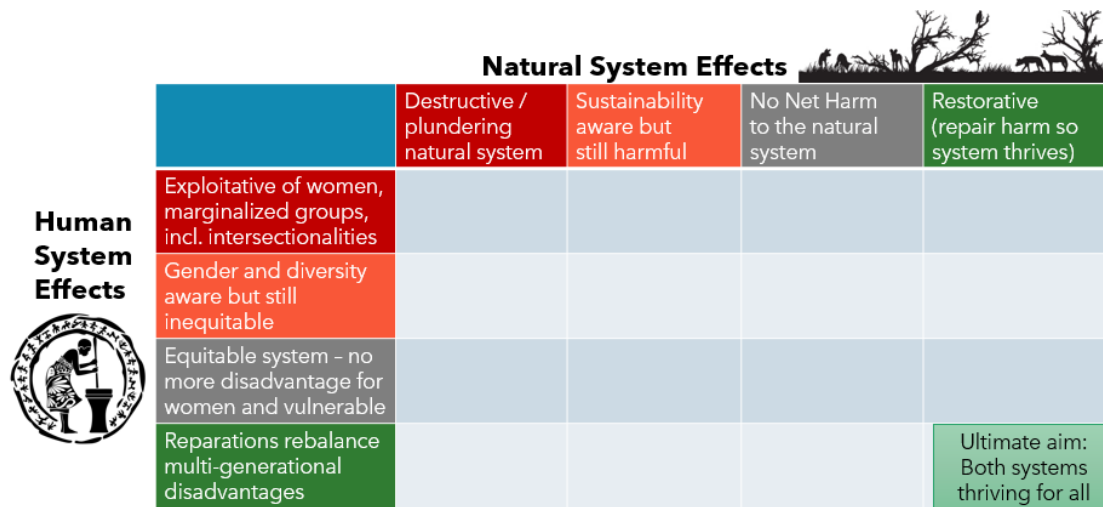
¹² [IPCC Sixth Assessment Report: Impacts, adaptation and vulnerability](#), 2022.

Expanding the typology to consider equity and sustainability together

GAC has been praised for its feminist evaluation methodology, which has helped bring into sharper focus how well women and other marginalized groups, including multiple intersectionalities, are served by various GoC policies and programs. However, the OECD-DAC peer learning review of mainstreaming environment in evaluation (2019) noted that this has unfortunately had the effect of pushing environmental concerns into the background.

How might evaluation teams move toward a more ambidextrous approach that retains its strong feminist lens while also bringing equally powerful focus to the environment? Another tool currently under development by Footprint Evaluation is the expansion of the 4-level typology into a 2-dimensional matrix that incorporates gender and equity concerns (Figure 9).

Figure 9. A 2-dimensional typology for assessing the equity and sustainability of policies and programs



		Natural System Effects			
		Destructive / plundering natural system	Sustainability aware but still harmful	No Net Harm to the natural system	Restorative (repair harm so system thrives)
Human System Effects	Exploitative of women, marginalized groups, incl. intersectionalities				
	Gender and diversity aware but still inequitable				
	Equitable system - no more disadvantage for women and vulnerable				
	Reparations rebalance multi-generational disadvantages				Ultimate aim: Both systems thriving for all

Source: Footprint Evaluation (forthcoming). Used with permission.

Again, the purpose here is to bring into sharper focus the reality of how beneficial or detrimental things are for women and other marginalized groups, as well as for the environment. This expanded typology could be used in several different ways. First, the 4-point “human system effects” scale could be applied when answering any evaluation question relating to design, implementation, outcomes, or impacts. Simply use the approach outlined above for the natural systems typology, drawing on multiple knowledges to ensure that your conclusions are robust.

A second approach would be to answer each evaluation question (or do an overall assessment across all questions) using the 2-dimensional matrix. This allows the evaluation to much more clearly convey how well the policy or program is addressing and meeting the needs of people and the environment, particularly the most fragile and damaged parts of the ecosystem and the groups that have historically been the most marginalized and/or harmed.

Finally, an assessment on the 2-dimensional matrix could be made for the policy or program overall, or for components of it separately.

Making clear evaluative judgements like this may seem risky or contentious for those new to the approach. This is why it is important to involve stakeholders, especially local and Indigenous knowledge holders, organisations working at the nexus between gender and the environment, and key decision-makers in the evaluative sensemaking process so that they participate in and fully understand the reasoning and the evidence behind each assessment.

It is not possible to involve everyone in the evaluative sensemaking process, and it is absolutely true that some receiving the report will inevitably disagree with the assessments. However, if the evidence and reasoning is laid out clearly and transparently, it is much more likely that the conclusions will be accepted as valid and credible. And if not, the clarity of the evidence and reasoning should allow for a fruitful discussion about areas of disagreement – for example, if some important effects or evidence are missing, or if the conclusions give too much or too little weight to certain effects or evidence, or if the reasoning contains a logical leap. The clearer and more user-friendly the report, the more easily a wide range of people can engage with it and help identify any issues with its evidence and reasoning, leading to a more robust evaluation report.

Those of us who use typologies or rubrics like this to draw clear, succinct evaluative conclusions have noted several advantages of this kind of approach.

First, the typology can be used up front in the project design stage, as well as in the evaluation design stage, to get stakeholders on the same page about what success needs to look like and what evidence would point to one conclusion vs. another, and importantly, to inform reflection whether the intervention should be modified before or early during implementation.

Second, when findings are presented as succinct evaluative conclusions, we have seen a dramatic increase in people's interest in how those conclusions were arrived at (compared with reports that are full of evidence but written without explicitly evaluative conclusions).

And finally, as we have already mentioned, clear evaluative conclusions are much better for communicating how serious the situation is and therefore how urgently decision-makers need to act.

The following resources will be useful to those who are new to the use of rubrics and explicitly evaluative reasoning:

- [What are rubrics?](#) [A short introduction (web page) from Jane Davidson.]
- [Actionable Evaluation Basics: Getting succinct answers to the most important questions](#) (Davidson, 2013; minibook). Also available in [Canadian!] French as [Les essentiels de l'évaluation tournée vers l'action: Obtenir des réponses succinctes aux questions les plus importantes](#) (Davidson, 2014; translated by Ghislain Arbour, with foreword by François Dumaine).
- [Evaluative reasoning](#) (Davidson, 2014. UNICEF Office of Research Methodological Briefs; also available in French as [Logique d'évaluation](#)). Evaluative reasoning is the process of synthesizing the answers to lower- and mid-level evaluation questions into defensible judgements that directly answer the key evaluation questions.
- [Evaluation methodology basics: The nuts and bolts of sound evaluation](#) (Davidson, 2005).

Step 11: Communicating findings and supporting use

Ensuring that evaluation is influential and used is no trivial task. That reality is magnified when incorporating environmental concerns into the evaluation of human system change efforts because some of our findings will be inconvenient truths.

Research shows that knowledge is more likely to be usable and used if it is salient to the decision-making processes in which it will be used; credible in the eyes of experts and those who will use the information; and seen as legitimate in a political sense.

To influence mindshifts, involve people in the evaluation process

When it comes to influencing decision makers who are primarily overseeing policies and initiatives to improve things in the human world (e.g., economic development, global security, education, and health) to make more environmentally sound decisions, we are seeking to influence their thinking in an enduring way, not just their short-term actions. A substantial shift in worldviews and mindsets is needed to help guide policies and other GoC initiatives that were not designed with the natural system in mind toward equitable and sustainable designs and ways of working that restore equity and enhance human wellbeing while *also* protecting and restoring the natural environment.

When we are looking for a deep and enduring shift in people's thinking, simply giving them useful information in a timely way isn't enough. Rather, we need to create opportunities for them to see the evidence with their own eyes and to puzzle with colleagues, other interests, and the evaluation team about its meanings and implications.

Identifying who we need to influence, and through whom

Let's start by clarifying who we are aiming to influence. During the inception phase of the evaluation, we already identified the primary intended users – the decision makers who will need the evaluation's insights to inform their decisions. It's likely that most or all of those identified are quite close to the policy or initiative being evaluated, but are they the ones with the authority to make the kinds of changes needed? Take another look at that list of primary intended users and consider whether anyone else needs to be added; for example, strategy and policy experts who write or influence policy and design programs and other policy instruments detailing how it will be enacted and enforced.

It's usually not possible to involve *all* of these people directly in the evaluation. A good, cost-effective alternative is to identify the key influencers among them. If we can involve those people directly in the evaluation process, they will be invaluable in helping us influence the others (Figure 10).

Thinking about the various people that the evaluation most needs to influence, especially the most important ones, ask the following:

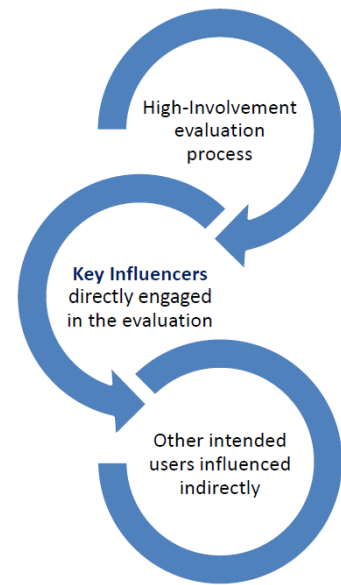
- Who do these people pay attention to? Who influences *them* when there's a significant change happening?
- Who could potentially *derail* the evaluation or its influence if they were *unconvinced* about the evaluation findings?
- Who would definitely need to be visibly involved in order to lend *political legitimacy* to the evaluation – either because of their positional power and influence or because they have a right to be involved?
- Whose or what *expertise* is needed to get the evaluation right, and for it to be *credible* in the eyes of various interests?

If we can involve the right key influencers *directly* in the evaluation process, they will be able to help us influence others *indirectly*.

Involving the right people at the right times and in the right ways

It is not feasible or worthwhile to involve the key influencers in all stages of the evaluation process. Instead, we need to figure out *who* to involve in *which parts* of the evaluation process and in *what ways*. As Figure 11 shows, there are many different parts of the evaluation where people might be involved. For suggestions about who to consider for each of these, please refer to the DFID UK source document.

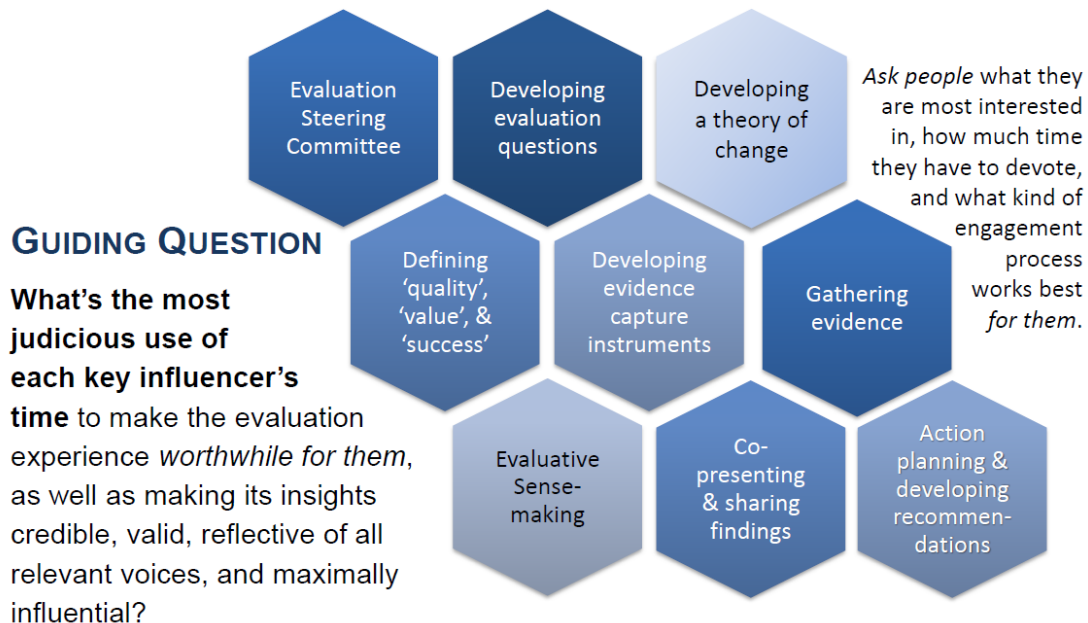
Figure 10. Involving key influencers in the evaluation is a cost-effective option



Source: Davidson (2017). *Influence by design*. DFID UK.

Figure 11. Key influencers should be involved in only those parts of the evaluation process that are interesting to them and to which they have something important to contribute

Source: Davidson (2017). *Influence by design*. DFID UK.



Is use and influence actually happening?

A critical step in ensuring use and influence is following up to check how well that is happening – and to do this early enough that the evaluation team could make adjustments to improve use and influence. This follow-up should go well beyond simply checking the management response to the evaluation to find out which recommendations were accepted.

One good option is having all members of the evaluation team, plus a selection of your key influencers, interview two or three primary intended users of the evaluation each. Interview a mix of the key influencers who were directly involved in the evaluation, as well as other primary intended users who were not directly involved.

The DFID guide includes questions to ask each of these groups, covering their experience being involved in and/or informed about the evaluation; how the evaluation is viewed by themselves and those around them; and the extent of actual use and influence (including conceptual, not just instrumental, use) so far. Be sure to include a question or two that probes how they now think about natural system effects and the coupled nature of human and natural systems. This will help gauge whether the evaluation is helping people see things in a coupled-systems light.

Resources:

- Davidson, E. J. (2017). [Influence by Design: How to scope, commission, and support influential evaluation](#). A user-friendly guide written for the UK Department for International Development; includes practical strategies and follow-up questions to ask primary intended users.
- Rowe, A., & Lee, K. (2012). [Linking Knowledge with Action: Promoting Use of Science Knowledge](#).

Call to Action

Interesting article [@Name]. We'd love to add your resource to the BetterEvaluation knowledge platform - if you'd be interested in sharing this with our users, please complete our resource contribution form and we'll take it from there. Thanks!

<https://www.betterevaluation.org/community/contribute-content/contribute-resource>

Bringing sustainability into evaluation faces two important challenges, both of which are fully within short reach of evaluation and evaluators to overcome. The challenges are largely within the politics, vision, habits and incentives for evaluation shaped on a worldview that environment does not merit consideration. Evaluation has the technical capacity to immediately address sustainability and important efforts are underway within evaluation to provide the tools, guidance and processes needed to progress this further¹³. The interest of GAC in adopting sustainability-inclusive approaches that led to this guidance are signature steps advancing the abilities of evaluation globally to address sustainability.

The first challenge we face is the urgency for interventions to cease harming natural systems and move as quickly as possible to restoring natural systems in ways that are also restorative of equity in human systems. Fortunately, some existing development interventions address coupled systems and seek restorative gains, for example in the IFAD evaluation referred to above almost a third of cases studies were achieving or exceeding do no net harm to natural systems and providing development gains for humans. All of these could be evaluated with existing knowledge from evaluation and natural sciences.

The second challenge is within the evaluation profession itself. Until recently, we have ignored natural and coupled systems, so that the collective evaluation experience, capacity and tools are lacking. This is improving and evaluators prioritising a sustainability-inclusive approach, employing mixed methods and focusing on conceptual use, will not be overly challenged by this. The key barriers are associated with worldviews that do not recognise the importance of natural systems and how sustainability and equity are coupled, with acceptance of decision processes and siloed approaches that work from a zero sum premise. These are only sometimes consciously-held views and are sometimes baked deep into organizational memory, policymaking, and initiative design. That is why evaluation needs to target and achieve conceptual use at this juncture, not just instrumental use (e.g., implementing recommendations).

GAC has two important capacities already in place that would otherwise have received more attention in this guidance. First, sustainability-inclusive evaluation benefits from collaborative approaches that reach beyond strict “stake in the intervention” interests. Feminist evaluation approaches already do this, for example reaching out to those with less power and beyond the strict accountability frames of interventions. Sustainability-inclusive evaluation adds more emphasis on human and non-human interests who can affect success or who are affected by the intervention. And GAC already employs mixed-method approaches. Secondly, the existence of a strong in-house environment division at GAC and the interest of both the evaluation and environment divisions in cooperating on sustainability-inclusive evaluation is a significant asset not widely present elsewhere.

¹³ Developing *sustainability-ready* evaluation capacities is a central mandate of Footprint Evaluation and includes tools, guidance and processes that will aid evaluation from a coupled sustainability and equity worldview. This guidance is an important element in establishing readiness and moving from guidance, tools and process to actually implementing *sustainability-inclusive* evaluations.

The authors of this guidance are enthused by the readiness of evaluation at Global Affairs Canada as global innovators in sustainability-inclusive evaluation positioned at the nexus of sustainability and equity.

Resources

Bruyninckx, H. (2022). New challenges for the evaluation community: Impactful contributions in times of urgency. *Evaluation*, 28(2), 144–149. <https://doi.org/10.1177/13563890221092950>

Rowe, A (2022) We can, we must: how should evaluation contribute to sustainability? in: What should evaluation learn from COP 26 ? Priorities of a cross-section of evaluation practitioners' Bour D, Brousselle A, Felcis W, Cekan J, Chaplowe S, Chelimsky E, Davies I, Leiter T, Menezes D, Picciotto R, Rogers P, Rowe A, Uitto J, Van den Berg R. in *Evaluation* 28 (1)

Rowe, A (2020) *Evaluation At the Nexus: Steps for Evaluating Sustainable Development Interventions*, in Juha I. Uitto (ed.) *Evaluating Environment in International Development: Contributing to National Results Beyond Projects*, Routledge second edition

Rowe, A. (2018) Sustainability-Ready Evaluation: A call to action, *New Directions Evaluation* 162 pp. 29-48 in George Julnes (ed.) *Evaluating Sustainability: Evaluation in Support of Managing Processes that Promote the Public Interest*

Rethinking Evaluation

During consultations with GAC, the environment division suggested a need to "rethink evaluation". This resonated strongly with us and our ambition for a "more urgent" and much more powerfully influential evaluation that is strongly connected to operational decisions about interventions, from their early conceptual phases through design, negotiation, and very early implementation.

The premise is simple: with only seven years to the forecasted era of irreversible thresholds in climate, biodiversity and other critical natural systems, and given that evaluations typically function primarily at longish intervals on a five-year cycle or longer, then when current initiatives are provided with evaluation advice they will be at or near the critical thresholds. And given that we are currently and widely still harming natural systems, then business-as-usual evaluation would be abdicating its responsibilities by sticking to the evaluation timeframes we have always used.

Footprint Evaluation also strongly advocates a rethinking of the evaluation function. Evaluation at GAC includes a Decentralised Evaluation Service Unit that provides support the shorter and earlier evaluation activities such as mid-term reviews or country evaluations. This seems to be where evaluation functions come close to operations. This is likely the lower hanging fruit for rethinking the evaluation function at GAC. And given the approaches employed by evaluation at GAC, the strong environment division within the Department and the relatively progressive development stance of the Department, lessons from rethinking evaluation at GAC could have global implications and possibilities.

To outline what this might look like, the evaluative thinking involved with Steps 1 to 3 provide the opportunity and vision of articulating likely sustainability and equity outcomes from proposed approaches: in Step 1 by including identification of likely outcomes on sustainability and equity, thereby addressing GoC and GAC commitments and policies; in Step 2 through opening the worldview to

recognise value of natural systems and coupling to equity in developing approaches; and in Step 3 identifying the nexus points for these approaches. Steps 8 and 9 are also important, taking a systems view of the proposed approaches and using this to gain insights about likely outcomes, using information such as from rapid evidence reviews, expert judgement and rapid evaluation assessments. A useful approach can be to connect this evaluative thinking to annual operational reviews.

Processes are an important part of the rethinking. The aim is conceptual use of the observations and implications from this rethinking approach; it is well established that joint knowledge production is an essential feature for such use. Thus, the evaluative processes need to be co-generated knowledge processes involving key interests in the initiative such as GAC program and operational areas, countries, and representatives of sustainability and equity interests. The rethinking approach starts at the design phase for initiatives and carries through inception to mid-term evaluation. With inception and early implementation, many realities occur such as delays in procurement, actual versus expected conditions, climate change and effects of unanticipated severe weather, and resulting shifting priorities of on-the-ground capacities. Interventions usually adapt, but without benefit of an evaluative rethink of the implications of options for sustainability and equity, applying an either-or mindset and often diminishing the importance and urgency of ceasing harm to natural systems. Rethinking evaluation in this way would be in effect truly applying the guidance on sustainability by illuminating the likely positive or adverse effects of adaptation options on environment and equity and prompting thinking about other win-win ways to adaptively manage interventions.

These processes and the knowledge generated would greatly benefit mid-term evaluations and empower them to substantively address evaluation criteria such as relevance and coherence and key evaluation questions about relevance and coherence, design and adaptation and implementation ([Step 5](#)). And importantly, evaluative thinking would be providing a valuable contribution towards capacity to design and adaptively manage interventions towards success in ending harm to natural systems and achieving equitable win-win solutions.