

VALUING VOICES
BETTER EX-POST EVALUATION WHITE PAPER SERIES

**PART 1: Toward Building Better Evaluation Practices: A Series on the Ex-
Post Evaluation Practices of Nordic Countries and the Netherlands: The
Search**

*By Preston Stewart, FDR Foundation/ Harvard Research Intern at Valuing Voices and
Jindra Čekan/ová, PhD Founder/ Catalyst at Valuing Voices at Cekan Consulting LLC*

June 2021

Keywords: Ex-post project evaluation, sustainability, Nordics (Denmark, Finland, Norway and Sweden) and the Netherlands,
methodology, data retention

Jindra@ValuingVoices.com

475 22nd Ave N. St Petersburg
FL 33704, USA
+1-202-848-4783

www.ValuingVoices.com

[@WhatWeValue](https://www.instagram.com/WhatWeValue)

Kolínská 5, Prague 130 00
Czech Republic
+420 727 980 535

Executive Summary/ Abstract

Ex-post completion or ex-post project evaluations are surprisingly rare given widespread commitments to ‘sustainable development’. Given the well-known commitment of the Nordic countries of Norway, Finland, Sweden and Denmark as well as the Netherlands, Valuing Voices, a consulting firm specializing in ex-post project evaluation, chose these countries as case studies. We hired an intern from Harvard College, who searched the national databases (in English only) and presented these results. The findings were surprising. There were far fewer ex-posts in total than expected. Those that were included a surprising number that was co-evaluated alongside final evaluations, although final evaluation illuminates relevance, effectiveness, efficiency (and the new OECD criteria of coherence), while ex-post project evaluation illuminates sustainability and impact two or more years after closure. There were lessons for Ministries about the searchability as well as the quality (and in the case of Denmark and in part, Sweden), the dearth of ex-posts at all. Overall, the recommendations for the five countries researched for this paper include better and more standardized definitions of ex-post project, greater absolute numbers done, transparent sharing of those done in publicly searchable manners, methodologically clear comparisons (baseline and midterm), and clarity in differentiating different evaluations.

There are caveats to these papers. This research was privately by Valuing Voices so the samples featured from the five countries were limited to public sources about ex-post and ex post evaluations. While we reached out to Ministry evaluative staff in all countries, one two made themselves available for consultation (the Netherlands and Finland) and provided a handful of additional ex-post evaluations. This paper focuses on what such research yielded, not definitive findings of programs or multi-year country strategies that are funded for 20-30 years continuously, nor projects funded by country-level embassies which did not feature on the Ministry site. We focus on project bilateral project evaluations, not multilateral funding of sectors. We also presented these findings in a [2021 webinar](#), during which we received input that Sweden’s EBA has a (non-project) portfolio of ‘country evaluations’ which looked back over 10 or even 20 year time horizons.. Still, we found very few ex-post evaluations at EBA (see cite in Search, below). No input was received by Norway or Denmark.

There are four papers in this series combined in this White Paper. First, the search across databases of all five Nordics (Norway, Finland, Sweden, and Denmark) and the Netherlands (pages 1-21). This is followed by country specific, detailed, papers for those with ex-posts: The Netherlands (pages 22-37), Norway (pages 38-47), and Finland (pages 48-59). We look forward to your feedback.

Jindra@ValuingVoices.com for the Valuing Voices team.

Research Process Overview:

In June-August 2020, supplemented by further research in October-November 2020, a search was conducted through government databases of the four Nordic countries and the Netherlands to identify ex-post evaluations of government-sponsored projects. This search through the evaluation practices of five countries: Nordics (Norway, Finland, Sweden, and Denmark) and the Netherlands took place over 50 hours (30 Stewart, 20 Cekan) . It should be noted that evaluations of financial investments in countries and regional settings were not included in the search if no specific project was attached. Furthermore, the search was only conducted in English, so there may be additional ex-post evaluations not published in English that are publicly available. The search was able to identify easily observable ex-post evaluations, but it is likely there exist other evaluations not turned up by the search, suggesting there is room for increasing the accessibility of such evaluations.

All five countries have written commitments to *sustainable development*¹ published on their own government websites and in development policy declarations.

Box 1: Nordics and the Netherlands Statements on Sustainable Development

Statements on sustainable development were found on the national development organizations' or central government websites and are as follows:

[From the MFA of the Netherlands](#) – “Through its development cooperation policy Dutch central government seeks to achieve the following goals in developing countries: ...Promote sustainable growth and climate action worldwide”

[From the government of Norway](#) – “The Government considers international efforts to bring about sustainable development to be of vital importance. Sustainable development also requires efforts at national level.”

[From the MFA of Finland](#) – “The 2030 Agenda [for Sustainable Development] includes 17 different goals that are related to sustainable development. These goals are to be met by 2030. Finland is committed to reaching them both at home and in its international cooperation. Finland’s development cooperation supports this commitment.”

[From the MFA of Sweden](#) – “A new development agenda was established in 2015 with the adoption of the 2030 Agenda, the Addis Ababa Action Agenda on Financing for Development and the Paris Agreement on climate change. For the first time, poverty reduction efforts in the international arena have been explicitly linked to all three dimensions of sustainable development – economic, social and environmental.”

[From the MFA of Denmark](#) – “Denmark can serve as a showcase for sustainable solutions. We can promote sustainable growth and employment in developing and growth countries and build upon our international experience and competencies.”

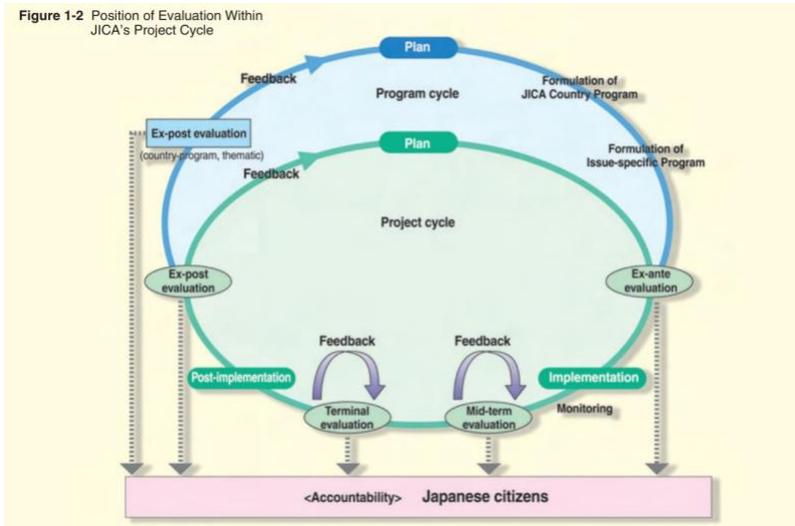
We at Valuing Voices believe that sustainable development can only be achieved when developmental aid donors return to measure the long-term sustainability of their projects at

¹ Valuing Voices attempted to contact evaluations teams from the five organizations researched for this paper. Out of the five organizations, only two responded – the MFA of Finland, and the Dutch IOB. Special thanks to the evaluations team from the MFA of Finland, specifically Ms. Anu Saxen and Ms. Sanna Pulkkinen, and Mr. Paul de Nooijer of the Dutch IOB for providing input on the paper series. Valuing Voices is interested in continuing our coordination with these organizations, as well as establishing contact with the other organizations studied and welcomes further input.

least two years after full exit from the recipient community has occurred by evaluating the durability of project outcomes and impacts, a process called “ex-post evaluation”. This definition is based on the [Japanese Development Agency’s \(JICA\) definition](#), which states:

“Ex-post evaluation is generally conducted... three years after the project[’s] completion with the emphasis on the effectiveness and sustainability of the project. This evaluation aims at deriving lessons learned and recommendations to improve the project as well as to help plan and implement more effective and efficient projects”.¹

The JICA definition is more precise but along the same lines of the 2019 [OECD definition of ‘sustainability’](#) as “the extent to which the net benefits of the intervention continue, or are likely to continue”.² Although the JICA definition states that three years (in general) pass before an ex-post evaluation is conducted, common practice among other organizations that conduct ex-post evaluations is to evaluate two or more (up to 20) years following project completion. We use the two to 15 year range at Valuing Voices to be past residual effects of donor inputs while also optimizing recall. These evaluations can help shape subsequent aid decisions of donor countries as well as how aid is designed and implemented for local sustainability. Our hope is for donor countries to invest in such ex-posts while carefully considering and learning from the findings of their ex-post evaluations.



*JICA’s project and evaluation cycle, which includes **accountability** to Japanese citizens.*

Image from:
https://openjicareport.jica.go.jp/pdf/11809803_02.pdf

Overall, the results from the search of ex-post evaluations in the national government ([Ministry of Foreign Affairs of the Netherlands](#) and [Ministry for Foreign Affairs of Finland](#)) and national non-profit (The Norwegian Agency for Development Cooperation ([NORAD](#)), Swedish International Development Cooperation Agency ([SIDA](#)), and Danish International Development Agency ([DANIDA](#)) databases of the Netherlands, Finland, Norway, Sweden, and Denmark were **not encouraging**. The global development projects for the latter three are posted on the NORAD,

SIDA, and DANIDA databases, respectively, which is why the search was not conducted on these countries' Ministry of Foreign Affairs websites. The databases were generally difficult to navigate, bringing up search results that were often unrelated to the keywords searched for (ex-post, ex post, impact evaluation) and returning untitled documents. With the term 'impact evaluation', the databases returned any documents that contained the term 'impact'. With the terms 'ex-post' and 'ex post', three different types of results were returned: purported ex-posts, calls for future ex-posts (which were counted and recorded) and mentions of the terms. The search revealed definitional problems, categorization problems, and database search problems.

Ex-post evaluations were rare. For instance, Sweden and Denmark did not publish any ex-post evaluations that were completed at least two years after aid ended. This does not mean that they were not conducted internally, only that none were publicly shared, unlike the three other donors who did share some of theirs. Overall, across all five countries, the term "ex-post" was often misused and the evaluation conducted during implementation or funding. There were also mis-categorizations, and ex-post evaluations were often published in reports containing final evaluations as well, leading confusion on which findings were ex-post, final, or mid-line.

Ultimately, the purpose of an ex-post evaluation is to shed light on the long-term *sustainability* of a project after project handover to local authorities, which is impossible to assess if the project is ongoing and funding is still coming from donors – the case in many of the purportedly 'ex-post' project evaluations that were actually final (endline) evaluations.

This paper is the first of a four-part series covering the search for ex-post evaluations in Nordic countries and the Netherlands. This paper covers the search process. Specifically, this paper will answer the following questions about each country before concluding with recommendations and findings from the search:

- Searched how: (where was searched, what keywords were used, how long did the search take, etc.)
- What was found: (including three specific subsections of findings: definitional findings, categorization findings, and search findings)
- What was done well regarding the country's ex-post evaluation practices
- What the ex-post evaluations found (sustainability-related findings will be covered in greater detail in the country-specific papers)
- What the organization can improve on in terms of its ex-post evaluation practices

The other three papers in this series will include more in-depth discussions of the ex-post evaluation practices found in 1) the MFA of the Netherlands, that had the greatest number of ex-posts but definitional issues, 2) in NORAD, which had many purported ex-posts but also definitional issues, and 3) in the MFA of Finland, which had the clearest definitions of terms but also only had four ex-posts. The purpose of these papers is not to tear down the good work of the five countries researched, but rather **to help build a culture of conducting ex-post evaluations that are clear in their purposes and hold to a single, agreed-upon definition of the term ‘ex-post’**. The criticism in this series of papers is intended to be constructive and contribute to a conversation about better evaluation practices.

The ex-post evaluations that were found were conducted on projects from various sectors, including from the agricultural/livestock (3), health (1), security (1), infrastructure (8), environmental (1), and education (2) sectors. Most projects emphasized gender equality as well.

Of particular note in the table below is the low number of actual ex-post evaluations and the dearth of calls for future ex-posts, considering all five countries’ commitments to sustainable development. Ex-post evaluation seems not to be a part of these countries’ evaluation norms.

Table 1: Summarized Definition/Categorization Search Findings -

	Search terms used / # of results per term	# of purported ex-posts	# of actual ex- posts	Calls for future ex-posts
The Netherlands²	ex-post (19 results), ex post (41 results), impact evaluation (160 results)	20	10	1
Norway	ex-post (701 results) impact evaluation (3987 results)	6	2	9
Finland³	ex post (132 results), impact evaluation (985 results)	3	4	5
Sweden⁴	ex-post (80 results), “impact evaluation” (5 results)	1	0	0
Denmark	“ex-post” (50 results), impact evaluation (71 results)	2	0	3
Totals		32	16	18

² The MFA of Finland and Mr. de Nooijer provided access to a [Finnish evaluation repository](#) and a [Dutch evaluation repository](#), respectively. However, neither of these repositories are divided by evaluation type, so they do not meet the criteria laid forth by this paper’s recommendations.

³ Two Finnish ex-posts and one Dutch ex-post are not summarized here because they were sent after the paper was initially written by representatives of the respective evaluation teams. Summaries may be found in the countries’ individual papers, issued separately by Valuing Voices. The norms for Finnish evaluation can be found [here](#). These norms do not mention ex-post evaluation, suggesting that Finland has yet to adopt ex-post evaluation as part of its evaluation process even though it has been defined.

⁴ The [Expertgruppen För Biståndsanalys \(EBA\) site](#) was highlighted as a potential source of ex-post evaluations on SIDA activities in our early 2021 webinar on ex-post evaluation practices. The EBA is a “government committee with a mandate to independently evaluate and analyse Sweden's international development assistance”, according to the [EBA website](#). . Out of a total of 60 publicly available reports only a few ex-posts were found, all of which were recently conducted, and the methods used in the ex-post evaluations were not standardized. Three of the ex-post evaluations read such as [the Evaluation of the Swedish Climate Change Initiative, 2009 – 2012](#) were quite weak methodologically, only glancingly visiting ultimate ‘beneficiaries’/ project participants. While this was a complex analysis of the Swedish Climate portfolio from 2009-19, the emphasis was on only one aspect of sustainability, further funding as an input. We encourage readers to go directly to the ex-post to look at the approaches taken and lessons learned for EBA, whom we nonetheless applaud for focus on learning from results post-completion.

The Search

1. The Netherlands

Searched how

Ex-post evaluations were searched for using the search bar on the [Policy and Evaluations Department \(IOB\)](#) webpage from the Ministry of Foreign Affairs of the Netherlands website. The search took six hours in total. The keywords used in the search and the number of results that were returned for each keyword are as follows: ex-post (19 results), ex post (41 results), impact evaluation (160 results).

What was found

Outside of a series of 10 infrastructure-related evaluations, some of which included ex-post evaluations or sub-evaluations, 10 purported ex-posts were found. Of those 10 purported ex-posts, three were actual ex-posts. Only one call for a future ex-post evaluation was found in the IOB's final evaluation documents.

The series of 10 evaluations focused on infrastructure evaluated 19 different projects in various countries that were active at some point in time between the years 2007-2012. Of the 19 projects that were evaluated, 11 were evaluated ex-post – the other eight projects were ongoing at the time of evaluation. Only seven of the 10 evaluations included ex-post evaluations of projects, yet they were all included in the [compendium](#) as if all were ex-post evaluations. In total, there were 20 purported ex-posts and 10 actual ex-posts from the Netherlands. One call for a future ex-post was found.

Definitional findings

Though impact evaluations are not necessarily ex-posts, in the case of the Netherlands, any document called an 'impact evaluation' was considered a purported ex-post due to the following quote from the IOB's "[Evaluation Policy and Guidelines](#)" publication (2009): "The RPE 2006 distinguishes between policy reviews and ex-post impact studies. Both concepts... are the equivalent of the more common terms 'policy evaluation' and 'impact evaluation'".³ The document states that the IOB treats the terms 'impact evaluation' and 'ex-post evaluation' as synonyms, so any impact evaluation is considered an ex-post by the IOB even though this is a misuse of the term 'ex-post'. While some of the impact evaluations are true ex-posts, seven of the impact evaluations are either midline or endline studies that *should not* be considered ex-posts by the IOB. Impact evaluations and ex-post evaluations are not the same. **All told, the IOB had difficulty defining evaluation-related terms.**

Categorization findings

Seven evaluations called ‘impact evaluations’ were found that were not ex-post evaluations even though the Dutch IOB referred to the two terms as synonymous. Only one ‘impact evaluation’ was found that was an actual ex-post evaluation, called “[Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice](#)”,⁴ published in the year 2012.

Search findings

Overall, the IOB database was difficult to navigate. Search results did not display how the term was used in the document. On other databases, a preview of each document showing the sentences the search term was found in allowed for faster, more efficient searches. Results were also not always relevant to the search terms, but fewer completely irrelevant results were returned on this database than on the other databases searched.

What did the IOB do well?

The IOB had the most ex-post evaluations of any of the five countries researched with a total of 10. On the database, documents were titled somewhat accurately, providing insight into what the documents contained.

*Findings of specific IOB ex-posts**

The Ixchen Centre for Women evaluation in Nicaragua (2012) found that cervical cancer screening rates increased during and after the 2005-2009 implementation of the intervention – the deployment of mobile cervical cancer screening services.⁴ 58% of the women who were given Pap smears by the Ixchen intervention sought out other sexual/reproductive health services after their cervical cancer screenings, indicating some success in raising awareness about the importance of taking care of one’s own SRH.⁴ However, the evaluators also found that the national government was hardly involved during aspects of project implementation, including the Pap screening activities. This was a problem for two reasons: first, that the government was a partner in this implementation but seemed shut out from aspects of it; second, that the government was supposed to take over the screening activities after donor ended implementation but were not associated with this intervention, causing women to wait for the mobile units to return (even though the mobile units were not going to come back).⁴ Greater national government involvement would have helped with the sustainability of the increased Pap smear coverage rates after the conclusion of the intervention.

The other ex-post that was not a part of the 10 infrastructure evaluations was conducted on a pair of projects called the [Reconstruction Programme \(RP\) and Strategic Partnerships in Chronic Crises Programme \(SPCC\)](#) that took place from 2012-2015 (RP) and 2014-2016 (SPCC).⁵ The evaluation was published online in 2019. The programs operated in 24 different countries, but the evaluation focused on findings from only three of those countries (Burundi, Ethiopia, and South Sudan). The RP and SPCC provided funding for many projects in each of those countries across different sectors.⁵ Though most of the findings were country-specific, one takeaway across all three was that violence created obstacles for project implementation.⁵ The evaluations found that the projects seemed poorly planned – and the programs had broad objectives that failed to guide the projects in any meaningful direction, especially during the RP.⁵ In a similar vein, they found that projects were poorly operationalized during the planning phase, superficially combining “various initiatives, such as local peacebuilding and socio-economic support and service provision”, but not meaningfully developing plans for those initiatives.⁵ Finally, the evaluators found that the fragmentation of human and financial resources hampered the sustainability of positive outcomes.⁵

[The series of 10 infrastructure evaluations](#) will be covered in greater detail in the Netherlands-specific post. However, there are two overall findings that should be highlighted. First, across the projects that were evaluated ex-post, it seemed like local governments were highly involved with the project implementations. Second, it seemed that sustainability concerns stemmed from concerns mostly about future funding for the interventions (such as governments pulling funding from continuing the interventions or taking the interventions in different directions than originally intended, although the latter can be a positive development).

What can the IOB improve?

The IOB should pin down the exact definition of what an ex-post is and separate the meanings of ‘impact evaluation’ and ‘ex-post’; they are not synonymous terms, as impact evaluations are not necessarily ex-posts. The “Evaluation Policy and Guidelines” should be edited to reflect a clearer acknowledgement of distinct definitions.

The IOB should work on providing document previews for the results of keyword searches.

Though there are a greater number of ex-posts written by the Netherlands than any other country researched, there should still be a stronger push to commission more ex-post evaluations.

The findings of the ex-post evaluations reveal weaknesses in the Netherlands' project planning. Project plans should offer specific guidance on how to work with local governments and set particular, quantifiable goals for their projects instead of offering broad guidelines. Careful project planning can be used to ensure local government involvement, foresee and address potential operationalization issues, and propose solutions for acting on post-project funding concerns.

2. Norway

Searched how

Ex-post evaluations were searched for using the search bar on NORAD website, which is the development cooperation branch of the Norwegian Ministry of Foreign Affairs. The search took three hours 45 minutes in total. The keywords used in the search and the number of results that were returned for each keyword are as follows: ex-post (701 results), ex post (701 results), impact evaluation (3987 results).

What was found

Six purported ex-posts were found, two of which were actual ex-posts. Nine calls for future ex-posts were found, but seven of those nine calls are for ex-posts were from a single initiative called REDD+ (a UN-led initiative). [A document](#) was also found referencing the Media Development Project in Mozambique that notes how an ex-post evaluation was proposed after the end of the project – a plan that was eventually scrapped.⁶

Definitional findings

Norway published contradictory definitions of the term 'ex-post'. According to a document called "[Evaluation of Norwegian Development Co-operation in the Fisheries Sector](#)" (2009), "Ex-post evaluations are carried out at regular intervals after the programme or project has started".⁷ Under this definition, any evaluation following the start of a project, including midline and endline evaluations, are incorrectly considered ex-posts. However, NORAD's documents do show some awareness of the accurate definition of 'ex-post'; according to an "[Organisational Review of the Norwegian Refugee Council \(NRC\)](#)" (2009), "there are also 'ex-post evaluations' taking place some years after a project or programme has finished in order to assess impact and sustainability".⁸ Though none of these NRC-produced documents are published on the NORAD website, the existence of this quote is an

acknowledgement that ex-post evaluations are not merely evaluations that occur after a project has begun.

Categorization findings

The two ex-posts found were not called ‘ex-posts’ in their titles. One of the evaluations was called an ‘impact evaluation’, while the other was called an ‘impact assessment’. Four documents were found calling themselves ‘ex-posts’ even though they were not – they were all either midline or endline evaluations.

Search findings

The NORAD database was the most difficult database to navigate out of the five databases used. Searching up the term ‘impact evaluation’ returned many irrelevant results after the first few search pages. The results a few pages into the search included any document with the word “impact” in it, even if it had nothing to do with evaluations.

What did NORAD do well?

The ex-post evaluations found were specifically commissioned so that the evaluations’ findings could be used to plan future projects (some of which have already begun as of the writing of this paper) taking place in the same regions and sectors as the projects that were evaluated. [One of the ex-posts](#), incorrectly called an ‘impact evaluation’, evaluated various development projects in West Pokot County, Kenya (2009),⁹ while the [other ex-post](#) evaluated NORAD’s support to the energy and road sectors in Zanzibar, Tanzania (2015).¹⁰ The intentionality behind these evaluations means that the findings have been or are going to be used for new projects in West Pokot and Zanzibar, showing that NORAD applies its ex-post evaluation findings to specific projects even though few ex-posts are available. These follow up projects were searched for on the NORAD database, but none were publicly available.

Findings of specific NORAD ex-posts?

The 2009 ex-post evaluation of development projects running from 1991-2007 in West Pokot, Kenya evaluated the sustainability of agricultural, sanitation, health projects carried out in interventions called the Pokot Development Project (PDP); its spiritual successor, the Pokot Integrated Project (PIP); and an HIV/AIDS training initiative geared toward training pastors and counsellors (prominent and visible community members) in teaching about HIV/AIDS prevention and treatment.⁹ According

to the evaluation, the “most successful and sustainable part of the development activities” was the construction of classrooms that are now used by local school committees to hold classes.⁹ Improved agricultural and sanitation practices were sustained, although not by the church that the intervention was turned over to.⁹ Instead, individuals in the area took it upon themselves to continue the improved practices.⁹ The main area where the intervention’s implementation failed was in educating women to gain literacy skills and knowledge of their human rights, largely because adult classes did not reach many women.⁹ Because of this implementation failure, the project was not able to sustain any gains in terms of women’s education outside of agricultural and sanitation training, which did reach women.⁹ Finally, the initiative to train pastors and counsellors (prominent and visible community members) in teaching about HIV/AIDS failed due to a lack of funding, so there were no positive outcomes to sustain.⁹ Unfortunately, little empirical data was collected for this evaluation, so all findings were taken from qualitative data.

The 2015 ex-post evaluation on Norway’s support to energy and road sectors in Zanzibar, Tanzania during 2004-2012 found that a high-capacity electric cable built by funding from Norway provided an “adequate and fairly reliable” electricity supply.¹⁰ Furthermore, the “road rehabilitation project” increased mobility.¹⁰ Combined with the improved electricity supply, these interventions led to greater access to health facilities and new business opportunities for economic diversification (such as the opening of stores selling fresh drinks and ice cream).¹⁰ However, electricity remained expensive, and although economic diversification was occurring in the area, at the time of the evaluation, there had been little poverty reduction.¹⁰ The evaluation also highlights that the project created some additional opportunities for women, but “many of the traditional constraints on the activity of women remain unaffected”, limiting the change possible for women.¹⁰

What can NORAD improve?

NORAD’s most notable issue – other than the dearth of ex-post evaluations publicly available – is the lack of clarity around the definition of the term ‘ex-post’. The contradicting definitions explain why two NORAD-sponsored midline and two endline evaluations are considered ex-posts. A specific webpage defining and collecting ex-post evaluations should be added to NORAD’s website.

NORAD should also double down on its efforts to promote and sustain gender empowerment. Though it is difficult to quickly change traditional values, as noted by the Zanzibar ex-post evaluation, the Pokot ex-post evaluation shows that even material, non-value-based efforts to advance the interests of women (such as through adult classes) have failed to produce satisfactory and sustained outcomes.

Steps should be taken in the project planning process to ensure that the intervention is implemented in a manner such that gender equality takes a prominent role during and after the intervention. NORAD cannot oversee local actors after an intervention has ended, but by setting up the groundwork for sustainable gender empowerment efforts during the intervention, NORAD can position local actors to continue the work started during the project.

Ex-post evaluations should also include quantitative data. The ease of producing quantitative data can be increased by **planning projects for ex-post evaluation during the project planning phase** so that project objectives can be set with quantitative data collection for future evaluations in mind.

3. Finland

Searched how

Ex-post evaluations were searched for using the search bar on the Ministry for Foreign Affairs of Finland website. The search took three hours 50 minutes in total. The keywords used in the search and the number of results that were returned for each keyword are as follows: ex-post (131 results), ex post (132 results), impact evaluation (985 results).

What was found

Five purported ex-posts were found, four of which were actual ex-posts. Five calls for future ex-posts were found in Finnish documents.

Definitional findings

The Ministry for Foreign Affairs of Finland published its [Evaluation Manual](#)¹¹ in 2018, which notes: Ex-post evaluation refers to an evaluation that is carried out some years after the programme has been completed. An ex-post evaluation provides evidence on the longer-term impact and sustainability of the programme. Ex-post evaluations are sometimes referred to as impact evaluations. Since impact is often included as an evaluation criterion also in other evaluations, the MFA does not use the term impact evaluation exclusively when referring to ex-post evaluation.¹¹

Of all the countries' databases found, this document has the strongest grasp on the definition of the term 'ex-post': a sustainability-focused evaluation of past projects intended to inform future developmental aid projects. Unlike the Netherlands and Norway, an inaccurate definition of the term 'ex-post' was not published anywhere on the Ministry for Foreign Affairs of Finland website.

Categorization findings

Though there were no inaccurate definitions of the term ‘ex-post’ in Finnish publications, there was a misclassification of an endline study as an ex-post evaluation.

Search findings

The Finnish database generally returned results relevant to the search terms – many of the documents the search turned up mentioned the keywords somewhere in the document.

What did the MFA of Finland do well?

Finland was the only country of the five that had ex-post evaluations *and* did not publish an inaccurate definition of the term ‘ex-post’.

*Findings of the MFA of specific Finland ex-posts**

The first of the Finnish ex-posts was on a project called [Strengthening of Environmental Administration at the Local Level in Nepal \(SEAM; 2001-2014\)](#) and was published in 2016.¹² Some outcomes of the project were positive, according to the evaluators: “Local level environmental administration is now established, strengthened and working with environmental plans and budgets included in the annual work plans”, and the project “contributed to and speeded up an ongoing process of decentralisation with regard to environmental administration and management”.¹² However, the evaluators noted that “there has been too little thinking and analyses of replication/policy/scaling up of models developed and introduced” to the national level;¹² the project did not outgrow its original goals. The evaluators noted that at the end of the project “there was no proper exit strategy and plan focusing on sustainability of results”, quashing hopes for scaling up the project and explaining why the outcomes had remained rooted at the local level two years after the end of the project.¹²

The other Finnish ex-post was on a project called the [Livestock Development Programme \(PRODEGA\)](#), which was based in Nicaragua and ran from 1990-2003.¹³ The ex-post evaluation was published in 2013.¹³ The project had a focus on promoting the dairy industry, and according to the evaluation, “10 years after the conclusion of the project, the survival of a part of the milk cooperatives created or supported by the project has been observed, along with the strong growth and strengthening of some of them. As a result, a large number of families have been able to improve their incomes and quality of life on a long-term basis”.¹³ The evaluation did find issues with the dairy industry that the project promoted, however. The evaluators note that the dairy industry is inefficient (“only 40% of

milk production is collected in the country”) and the milk is of low quality – both of which are problems attributed to a need to maintain quality standards in the Nicaraguan dairy industry.¹³

What can the MFA of Finland improve?

Although the definition of the term ‘ex-post’ found in the Ministry’s documents is strong, an endline study was called an ex-post evaluation by the evaluators. This misclassification serves as a reminder to circulate the definition to assure that evaluators are aware of the term’s meaning.

From the findings of the ex-posts, it seems the Finnish projects could be improved by planning ways for the projects to expand as time goes along following Finland’s exit. Both the institutions put into place by the SEAM and PRODEGA interventions have struggled to outgrow their original designs – post-SEAM, the local climate change administrative gains have not been scalable; post-PRODEGA, the dairy industry has had difficulty modernizing. Neither of these issues fall squarely into Finland’s court – they have not been with either intervention for multiple years at this point – but exit strategies could have been developed that would have helped the actors who took over the projects to expand the scope of what the intervention accomplished.

4. Sweden and 5. Denmark

Searched how

Ex-post evaluations were searched for using the search bar on the SIDA (Sweden’s development agency) and DANIDA (the term for Denmark’s development efforts) websites. The search took two hours 30 minutes in total for the Sweden search and one hour 15 minutes total for the Denmark search. For Sweden, the keywords used and the number of results that were returned for each keyword are as follows: ex-post (80 results), ex post (80 results), “impact evaluation” (5 results). For Denmark, the keywords used in the search and the number of results that were returned for each keyword are as follows: “ex-post” (50 results), “ex post” (50 results), impact evaluation (71 results).

What was found

One purported ex-post evaluation from Sweden and two purported ex-posts from Denmark were found. However, none of these were actual ex-posts*. No calls for future ex-posts from Sweden were found, but there were two from Denmark.

Definitional findings

Unfortunately, no definitions of the term ‘ex-post’ were published by either country.

Categorization findings

The [purported ex-post from Sweden](#), evaluating health sector projects in Russia, stated that it was a combination of ex-post evaluation and ongoing project evaluation:

“This evaluation assesses the support SIDA provides to the social and health sector in North West Russia. The evaluation is a combination of an on-going and an ex-post evaluation. The evaluation has a simultaneous focus on the implementation process, as well as the results and impacts of twenty four health and social sector projects, spanning a wide cross section of support in North West Russia”.¹⁴ It was difficult to tell which parts of the report were ex-post and which parts were ongoing project evaluation, so it was not classified as an actual ex-post.

One of Denmark’s purported ex-posts that evaluated Danish aid in Uganda from 1987-2005 was unclear about which projects were ongoing by 2005 and which projects had been completed.¹⁵ Denmark’s other purported ex-post evaluated ongoing projects: “This was an ex-post evaluation of the Fast Start Climate Change (FSCC) commitments made by Denmark until the end of 2013. For some projects implementation is still ongoing”.¹⁶

Search findings

Neither country’s databases were too difficult to navigate, but results were largely irrelevant (considering how few ex-posts were found).

What did SIDA and DANIDA do well?

In all likelihood, the two miscategorized purported ex-post evaluations (one from each country) did contain some ex-post evaluation elements. The inclusion of ex-post elements in these evaluations suggest that ex-post evaluations have been attempted by Sweden and Denmark. Nonetheless, but evaluations of what is sustained cannot be performed during or even immediately after implementation.

What can SIDA and DANIDA improve?

Both countries should commit to conducting ex-post evaluations and publishing findings in a document or online statement. In the same document or statement that calls for ex-post evaluations, both countries should clearly define their expectations for what an ex-post evaluation should be so that evaluation consultants have explicit guidelines to follow. A specific webpage on the SIDA and DANIDA sites should also be marked as a designated location for the posting of ex-post evaluations.

General Recommendations

Overall, the recommendations for the five countries researched for this paper are to:

1) firmly define what an ex-post evaluation is and hold close to this definition and

2) conduct a greater number of ex-post evaluations under the new definitions, committing to ex-post evaluation as a norm in institutional M&E practices. The low total of ex-post evaluations found, as well as the dearth of calls for future ex-post evaluation, suggest that these countries do not yet regularly conduct ex-post evaluations.

3) This paper also calls upon the developmental organizations to publicly release ex-post evaluations in clearly marked locations on their websites to increase accountability and accessibility. JICA, for example, has [a webpage dedicated specifically to publishing its ex-post evaluations](#).¹⁷ Such publicly available evaluations should become the norm in the Nordic countries and the Netherlands. Hopefully, the OECD will create a by-country repository offering access to ex-post evaluations at some point in the future. Furthermore,

4) baseline statistics should be measured using quantitative indicators decided upon at the outset of interventions, and such indicators should be monitored throughout the intervention. This way, ex-post evaluations can measure the same indicators on return trips to the project sites and claims of success or failure can be quantified.

Finally,

5) though a logistical challenge that may be difficult to solve due to funding restraints, evaluation reports should ideally only contain a single evaluation type – or at least clearly mark off ex-post sections from endline sections, mid-line from endline, etc. Even considering funding constraints, evaluation reports should be organized to prevent confusion.

This research has led to **three additional individual case-study** documents: the Netherlands, Finland and Norway. Valuing Voices plans to publish a white paper on this research in 2021. Valuing Voices hopes the recommendations will be adopted so that ex-post evaluations become the norm in evaluations processes. Ultimately, these recommendations are intended to increase the capacity for transparent learning from previous projects as well as catalyze interest in learning from what was and was not sustained, why, and how we can change our project cycle to foster sustained results. We strongly encourage Ministries in the Netherlands and the Nordic countries featured to take these results and invest in evaluations of sustainability and databases.

Research note on the search process:

The keywords chosen (ex post, ex-post, impact evaluation), variations of which were used across all five of the countries researched, were chosen because they relate to ex-posts; “impact evaluation” was chosen because of the common practice of labelling ex-posts as ‘impact evaluations’. The terms ‘sustained’ and ‘post completion’ were searched following the initial search with the three aforementioned terms, but these keywords did not bring up any other ex-posts. The number of purported ex-posts found were counted, along with how many of those ex-posts actually fit the definition of ‘ex-post evaluation’ for each country. As numerous documents found in the search emerged because they asked for future ex-posts to be carried out on the existing final evaluation, how many documents each country published calling for ex-post evaluations to be carried out in the future was also counted.

Definitional note on the term ‘impact evaluation’:

This paper uses the term ‘impact evaluation’ as it is defined by the OECD in [this document](#), which states that an “impact evaluation is an assessment of how the intervention being evaluated affects outcomes, whether these effects are intended or unintended”.¹⁸

Search Paper Citations

1. JICA. (n.d.). Ex-post Evaluation (Technical Cooperation). Retrieved October 1, 2020, from https://www.jica.go.jp/english/our_work/evaluation/tech_and_grant/project/ex_post/index.html
2. OECD/DAC Network on Development Evaluation. (n.d.). Better Criteria for Better Evaluation. Retrieved January 4, 2021, from <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>
3. IOB. (2009, October). Evaluation policy and guidelines. Retrieved June 13, 2020, from <https://english.iob-evaluatie.nl/publications/regulations/2009/10/01/evaluation-policy-and-guidelines>
4. IOB. (2012, October). Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice . Retrieved June 13, 2020, from <https://english.iob-evaluatie.nl/publications/evaluations/2012/10/01/371---impact-evaluation-of-the-ixchen-centre-for-women-cervical-cancer-programme-in-nicaragua-2005-2009---turning-a-right-into-practice-kopie>
5. IOB. (2019, July 1). Less Pretension, More Realism: An evaluation of the Reconstruction Programme (2012 - 2015), the Strategic Partnerships in Chronic Crises Programme (2014 - 2016) and the Addressing Root Causes Tender Process. Retrieved June 8, 2020, from <https://english.iob-evaluatie.nl/publications/evaluations/2019/07/01/428---iob---an-evaluation-of-the-reconstruction-programme-2012-2015-the-strategic-partnerships-chronic-crises-programme-2014-2016-and-the-addressing-root-causes-tender-process>
6. Disch, A., & Taela, K. (2010). Democracy Support through the United Nations. Retrieved July 13, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/vedlegg-til-publikasjoner/case-report-moz1.pdf>
7. MRAG. (2009, January). Evaluation of Norwegian Development Co-operation in the Fisheries Sector. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/vedlegg-til-publikasjoner/evaluation-of-norwegian-development-co-operation-in-the-fisheries-sector1.pdf>
8. Bain, J., & Sørnum, V. (2009, September). Organisational Review of the Norwegian Refugee Council. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/vedlegg-til-publikasjoner/organisational-review-of-the-norwegian-refugee-council.pdf>
9. Kristensen, K., & Nairesiae, E. (2009, October). IMPACT EVALUATION OF THREE PROJECTS IN POKOT, KENYA. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/ngo-evaluations/impact-evaluation-of-three-projects-in-pokot-kenya.pdf>
10. The International Law and Policy Institute. (2015, March). Impact assessment of Norwegian support to the energy and road sectors in Zanzibar. Retrieved June 13, 2020, from <https://NORAD.no/globalassets/publikasjoner/publikasjoner-2015-/NORAD-collected-reviews/impact-assessment-of-norwegian-support-to-the-energy-and-road-sectors-in-pemba.pdf>
11. MFA of Finland. (2020, June 16). Evaluation Manual. Retrieved January 4, 2021, from <https://eoppiva.zapter.io/evaluationmanual2018>

12. Axberg, G. N. et al (2016, June 30). Post-Evaluation of Strengthening of Environmental Administration at the Local Level in Nepal (SEAM), Years 2001 – 2014. Retrieved June 9, 2020, from https://um.fi/documents/384998/0/Nepal_SEAM_Post_evaluation%2BTOR_2016.pdf/03bc624c-2241-7fd0-2c03-0d9b3c68feda#search=%22ex+post%22
13. Seppänen, M., Eriksson, L., Aguilar, B., Boman, K., & Pijnenburg, T. (2013, May). Evaluation of Rural Sector Cooperation between Nicaragua and Finland. Retrieved June 9, 2020, from https://um.fi/documents/384998/385866/prodega_final_report/b4812a1d-5bc0-b467-3634-4fd13763cb8e?t=1528278516971
14. Bjørnkilde, T., & Wynn, A. (2004, April). Social and Health Sector Projects in Russia. Retrieved June 8, 2020, from <https://www.SIDA.se/contentassets/88db8b0f4910498a84845316ecdd4d5d/14282.pdf>
15. DANIDA. (2006, November). Evaluation of Danish Aid to Uganda 1987–2005. Retrieved June 8, 2020, from https://um.dk/en/DANIDA-en/results/eval/eval_reports/evaluations/publicationdisplaypage/~media/um/english-site/documents/DANIDA/eval/87-7667-626-9/200606ugandavol4thematicpapers.pdf
16. DANIDA. (n.d.). Country (Kenya) Sub-Evaluation Executive Summary. Retrieved July 6, 2020, from https://um.dk/en/DANIDA-en/results/eval/eval_reports/evaluations/publicationdisplaypage/~media/um/danish-site/documents/DANIDA/resultater/eval/201504annex6.pdf
17. JICA. (n.d.). Ex-post Evaluation (ODA Loan). Retrieved July 28, 2020, from https://www.jica.go.jp/english/our_work/evaluation/oda_loan/post/index.html
18. OECD/DAC Network on Development Evaluation. (2019, November/December). OUTLINE OF PRINCIPLES OF IMPACT EVALUATION. Retrieved July 13, 2020, from <http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf>

VALUING VOICES

BETTER EX-POST EVALUATION WHITE PAPER SERIES

PART 2: Toward Building Better Evaluation Practices: A Series on the Ex-Post Evaluation Practices of Nordic Countries –The Netherlands

By Preston Stewart, FDR Foundation/Harvard Research Intern at Valuing Voices

This paper will cover government sponsored ex-post evaluations from the Netherlands. As noted in the first paper of the series, the key findings for the Netherlands were as follows:

- The Policy and Operations Evaluation Department (IOB) of the Ministry of Foreign Affairs of the Netherlands (MFA) had the greatest number of ex-post evaluations of any Nordic country with a total of 10 found in its full collection. However, the vast majority of those evaluations came from a single commissioning of 10 evaluations of infrastructure projects in seven different countries (Bangladesh, Ghana, Sudan, Guatemala, Indonesia, Sri Lanka, and Tanzania) evaluated on project activities between 2007-2012. The series was published online in 2015, and of the 19 projects featured, 11 projects were evaluated ex-post (across seven evaluations) and eight were final or midterm. The remaining three ex-post evaluations were of projects in three African countries (Burundi, Ethiopia, and South Sudan) Nicaragua, and Afghanistan, respectively.
- Yet overall, there was difficulty pinning down a definition of the term ‘ex-post’. The IOB published an “[Evaluation Policy and Guidelines](#)” document calling the terms ‘ex-post’ and ‘impact evaluation’ synonymous.¹ However, impact evaluations and ex-post evaluations are not the same according to the [OECD definition](#) of the term ‘impact evaluation’, which states that an “impact evaluation is an assessment of how the intervention being evaluated affects outcomes, whether these effects are intended or unintended”.² The search turned up seven documents called ‘impact evaluations’ that were not ex-post evaluations.
- Only one document was found calling for an ex-post evaluation to be conducted in the future.

What are ex-post evaluations? [According to the Japan International Cooperation Agency \(JICA\)](#), the global leader in conducting ex-post evaluations, ex-post evaluations are generally conducted years after project completion and aim at “deriving lessons learned and recommendations to improve the project as well as to help plan and implement more effective and efficient projects”.³ This series of analyses of ex-post evaluations aims in part to spark a conversation about definitional accuracy within the Nordic development community.

Definitional Findings

While the ex-posts offer a plethora of lessons about what was sustained (or not sustained) and why, one overarching finding was a base-misunderstanding of how the term ‘ex-post’ is defined and how ex-post evaluations are conducted. The IOB considers the terms ‘impact evaluation’ and ‘ex-post evaluation’ synonymous (even though they are not), as evinced by the IOB’s “Evaluation Policy and Guidelines” publication: “The RPE 2006 distinguishes between policy reviews and ex-post impact studies. Both concepts... are the equivalent of the more common terms ‘policy evaluation’ and ‘impact evaluation’”.¹ However, impact evaluations are often conducted during project implementation*. On the other hand, ex-post evaluations can *only* occur after the end of a project and evaluate *sustainability*. The “Evaluation Policy and Guidelines” display a lack of clarity in defining the term ‘ex-post’.

Categorization findings

Improvements could also be made to the classification of the actual Dutch ex-posts. The IOB has called ex-post evaluations, as well as endline and midline evaluations, ‘impact evaluations’. **All but one of the actual ex-posts found were not referred to as ex-post evaluations in their titles:**

- One of the ex-posts is called “[Evaluation of the Reconstruction Programme, the Strategic Partnerships Chronic Crises Programme and the Addressing Root Causes Tender Process – Less Pretension, More Realism](#)” (ended in 2016, evaluated in 2019).⁴
- Another is called “[Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice](#)” (ended in 2009, evaluated in 2012).⁵
- The set of infrastructure project evaluations is called “[Country case studies – Investing in Public Infrastructure](#)” (ended in various years, up to 2012 for projects evaluated ex-post, evaluated in 2014-2015).⁶

Furthermore, out of eight documents called “impact evaluations”, only one of them was an actual ex-post evaluation (“Impact Evaluation of the Ixchen Centre...”).

It is important for the titles of evaluations to reflect evaluation type. If these evaluations had ‘ex-post’ in each of their titles and the IOB provided a clear, accurate definition of the term ‘ex-post’, readers would immediately understand that these reports contain post-project evaluations discussing the sustainability of outcomes. It is understood that various types of evaluations are grouped, which may make classification somewhat difficult, but accurate titling should nonetheless be attempted on evaluations that clearly fall into categories. Ideally, though a logistical challenge that may be difficult to solve due to funding restraints, evaluation reports should only contain a single evaluation type – or at least clearly mark off ex-post sections from endline sections, mid-line from endline, etc.

Search findings

The IOB database was difficult to navigate. The other Nordic databases contained previews of where the search terms were mentioned in each document. The search of the IOB database only yielded document titles that hinted at what was within the document – as there was no context cited, the search was made more difficult. On a positive note, document titles were somewhat informative (given the definitional issues above), unlike other databases that returned documents titled with numbers. In general, results were fairly relevant to the search terms inputted. Though improvements to the database could be made, a more useful tool than a database could be created: an online repository of ex-post evaluations that is easily accessible via a link on the IOB website, akin to [this one](#) by JICA.^{7*}

Ex-Post Evaluations Found

The IOB published three ex-post evaluations on their website that were not from the series of 10 infrastructure project evaluations. This section includes a list of the ex-post evaluations found. Summaries of the findings from the ex-post evaluation reports related to the *sustainability* of results can be found in Annex 1.

1. [Evaluation of the Reconstruction Programme, the Strategic Partnerships Chronic Crises Programme and the Addressing Root Causes Tender Process – Less Pretension, More Realism](#)⁴
2. [Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice](#)⁵

3. [Post-mission evaluation of the integrated police training mission in Kunduz](#)^{8*}

From the ORET evaluation series:

4. [Railway Signaling and Interlocking in Bangladesh](#) (ended 2005)⁹
5. [Dutch Buses for Ghana](#) (ended in 2010)¹⁰
6. [Drinking Water Infrastructure Projects in Ghana and Sudan](#) (ended in 2013)¹¹
7. [One Buoy Tender and Three Aid Vessels to Indonesia](#) (ended in 2009)¹²
8. [Upgrading the Disaster Response Network in the Western Province and Selected Urban Areas in Sri Lanka](#) (ended in 2012)¹³
9. [Airport Rehabilitation in Tanzania](#) (ended in 2012)¹⁴
10. [Rehabilitating Diagnostic Services in Tanzania](#) (ended in 2006)¹⁵

Recommendations

The Dutch MFA should properly classify their reports according to the accurate definition of evaluation types. If an evaluation is actually an ex-post, it should be called an ex-post evaluation in its title. If an evaluation is an impact evaluation, evaluating the effectiveness of an array of intervention activities during implementation, its title should refer to it as an impact evaluation. **The MFA would benefit from clear guidance highlighting a commitment to careful definitions and should create a centralized online location to access ex-post reports focused on sustainability, differentiated from other kinds of evaluations.** This way, it would be easier to learn from differential evaluations so that the findings can yield clear lessons to be used for future projects.

Additionally, the findings from the specific ex-post evaluations (summarized in Annex 1) have led to the following recommendations:

- Though the ORET evaluation series evaluated projects that closely involved relevant local and national governments, the Ixchen intervention failed to plan inclusion of the Nicaraguan government adequately. **Valuing Voices' recommendation is for the project planning phase to specifically include strategies for partnerships between relevant government and implementation actors.** In the Ixchen intervention, for example, the project plan should have included a specific plan for integrating the Nicaraguan Ministry of Health into Pap smear education campaigns. Nicaragua has conservative policies on women's health that could have posed challenges for close coordination, but the intervention highlighted the government as a key entity for the intervention.⁵ The intervention should either have planned for another group

to take over or attempted to loop in government workers as much as possible (since, as described in Annex 1 and the evaluation report, part of the problem was the physical absence of MINSA workers during intervention activities).⁵ Similarly, the Tanzanian government's capacity to sustain the health equipment was overestimated. The new medical equipment sent to Tanzanian hospitals fell into disrepair due to "irregularities in the... flow of funds from government to the facilities".¹⁵ **The sustainability post-handover costs must be considered and planned for jointly, and more broadly, general exit strategies should be formed to increase the likelihood of sustainability.**

- The ORET evaluations and RP/SPCC evaluation highlight **the importance of considering country context before pushing an intervention forward. Sustainability should be a priority for interventions from the outset of the interventions' planning phases.** In the case of some RP/SPCC projects, violence hampered the implementation of projects to the point where sustainability became impossible. A number of the ORET interventions were implemented successfully, but then were not sustained because of other government priorities. The Tanzanian airport rehabilitation and diagnostic rehabilitation cases, as well as the Sri Lankan fire brigade case, show how **government priorities can affect the sustainability of interventions.** Changes made to the airport were well-sustained because the Tanzanian government was enthusiastic about maintaining the intervention's improvements. Though the MFA of the Netherlands cannot be blamed for contextual circumstances or the failure of local governments to sustain interventions, **Valuing Voices recommends that the MFA considers the prospects that local priorities and contexts lend themselves to regarding project sustainability when deciding where to allocate resources.** This does not mean refraining from interventions in areas with violence and general instability; however, this recommendation suggests that **sustainability should be factored into decisions about where to conduct interventions** to avoid projects that ultimately lead to few positive outcomes. A greater emphasis on conducting and learning from ex-post evaluation would help the MFA identify factors contributing to outcome sustainability and better inform decision-makers on what projects may succeed. Of course, this does not mean every project will lead to sustained positive outcomes; local contexts change, sometimes rapidly, and instability will always be a concern in development work.
- The RP/SPCC evaluators found that projects' goals were often too ambitious or only superficially attempted to address to various problems (such as gender inequality), leading to poor implementation and sustainability of positive outcomes. These problems are signs that the programs' broad goals did not offer enough guidance to organizations planning projects

under these programs. **Valuing Voices recommends that programs and projects lay out clear, measurable outcome and impact goals during the planning phase for implementation organizations to follow and which can be evaluated ex-post.** Calls to action such as the RP's goal of "improving human security" offer little guidance in the way of how to plan or measure specific, targeted interventions.

- Finally, it was found that some of the evaluations, particularly the ORET evaluations, were lacking in quantitative data. **Valuing Voices believes baseline statistics should be measured using quantitative indicators decided upon at the outset of interventions, and such indicators should be monitored throughout the intervention.** This way, ex-post evaluations can measure the same indicators on return trips to the project sites and claims of success or failure can be quantified for transparent learning.

Endnotes for the Netherlands

- Special thanks to Mr. Paul de Noojer, a coordination policy researcher from the IOB, for lending his input on this paper. Mr. de Noojer provided the author with the "Post-mission evaluation of the integrated police training mission in Kunduz" ex-post evaluation (the full report of which may be [found here](#) in Dutch) and the repository of IOB evaluations, as well as clarification on various points included in the paper. The paper has been edited to reflect his suggestions.
- Impact, [according to the OECD](#), assesses the effects of the intervention, as opposed to sustainability, which assesses whether benefits effects will last. Though some impacts are not apparent during project evaluation, impact evaluations can be (and often are) conducted during interventions to assess immediately visible impacts to guide the rest of the intervention.
- Mr. de Noojer provided access to an [IOB evaluation repository](#) that lists all evaluations conducted by the IOB. However, this repository is not grouped by evaluation type, unlike the JICA repository that is mentioned as a model.
- The Netherlands' commitment to sustainable development: [From the MFA of the Netherlands](#) – "Through its development cooperation policy Dutch central government seeks to achieve the following goals in developing countries: ... Promote sustainable growth and climate action worldwide"

Annex 1: Cited Project Details

1. The ex-post called “Evaluation of the Reconstruction Programme, the Strategic Partnerships Chronic Crises Programme and the Addressing Root Causes Tender Process – Less Pretension, More Realism” focuses on the Reconstruction Programme (RP) (2012-2015), which overlapped with the Strategic Partnerships Chronic Crises Programme’s (SPCC) (2014-2016) activities in three countries: Burundi, Ethiopia, and South Sudan.⁴ This ex-post evaluation was published online in July 2019. Both programs were intended to promote projects that addressed the “root causes of poverty, conflict, terrorism, climate change, population growth and irregular migration” to foster “stability and self-reliance in fragile contexts”.⁴
 - The project objectives for the Reconstruction Programme broadly called for Reconstruction Programme projects to improve “human security” (including “economic security, food security, health security, environmental security, political security, personal security and community security”) and increase the “peace dividend” (defined by the evaluation as “the benefits gained from reduced military spending and income gains after the conflict ends”).⁴
 - The SPCC objectives were more specific, focusing on “security and the rule of law”, food security, and addressing population flows and migrant and/or refugee needs.⁴ The evaluation found that projects from both programs struggled to create lasting impact after the end of implementation.⁴ Furthermore, the projects’ implementations were largely hampered by the contexts in which they were conducted; oftentimes, it was difficult to implement projects in areas affected by violence.⁴ The evaluation recommended that the MFA of the Netherlands should focus developmental programming on addressing conflict.⁴ The MFA has already taken some action on this front; in 2014, [the MFA of the Netherlands established a fund](#) of 570 million Euros “to meet the most immediate needs of those affected by conflict-related violence and natural disasters”.¹⁶ It remains to be seen whether the MFA will make additional contributions to violence reduction in light of the findings of this ex-post evaluation.
 - Projects were sometimes too ambitious and often superficially combined “various initiatives, such as local peacebuilding and socio-economic support and service provision”.⁴ Furthermore, according to the ex-post, “the achievement and sustainability of results have been limited by high levels of fragmentation, as resources were spread too thinly between numerous small and geographically dispersed activities”.⁴ The resources that were spread too thin included financial and human resources – there were shortages of adequately trained staff on many projects.⁴ Overambition during the planning process and fragmentation during implementation resulted in the poor sustainability of positive outcomes.

2. The other ex-post that did not originate from the series of 10 infrastructure project evaluations was called “Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice”.⁵ The objective of the project was to improve knowledge of, as well as provide screenings for, cervical cancer through a partnership between the Nicaraguan Ministry of Health and the local NGO ‘Ixchen’.⁵ The mobile screening units project ran from 2005-2009, and the evaluation was conducted in October 2012.⁵

Positive findings:

- The evaluation found that the project had strong results during implementation and relatively so after closure. The local sustainability of demand for health services was strong – the mobile units of the local NGO that the intervention partnered with were able to raise the rates of cervical cancer screenings during and following the end of the intervention.
 - Pap smear coverage rates rose from 7-11% of women aged 14-45 in 2003-2005 (before the Ixchen intervention) to 33.76% of women during the project, then fell to 21-26% of women in 2009-2011 after the Ixchen intervention.⁵ These statistics indicate that a positive trend in coverage rates was sustained after the intervention, although coverage rates were not as high as during the intervention.
 - “Following the care by Ixchen, one of the results sought was to create long-term awareness among clients regarding the services, which would be reflected in the use of other SRH [sexual and reproductive health] services, the channeling of demand towards other health facilities and the reorientation of the women in the search of these types of services, which are normally not seen as a necessity. Of all the women surveyed, 58% sought other SRH services [three years] after the Ixchen intervention”.⁵

Negative findings:

- The most notable failure of the intervention was the lack of public-private partnership between the government and the local non-governmental organization and health ministry in charge of the Ixchen intervention. The evaluation states: “the failure to involve MINSAs [the Nicaraguan Ministry of Health] personnel actively in adopting the rights-based approach [of the Ixchen implementation] and taking part in educational and Pap screening activities was a missed opportunity for transferring skills [between the Ixchen workers and MINSAs workers]. Programme coordination at national level was virtually non-existent”.⁵ MINSAs staff were largely relegated to complimentary roles during the educational and Pap screening activities

instead of being given the primary roles they should have had.⁵ As a result, some women (no exact number was specified) did not associate the government with the provision of cervical cancer screenings and did not seek out Pap smears from MINSA clinics following the end of the intervention even though screenings were available.⁵

- Though the aforementioned failures occurred during project implementation, they have far-reaching impacts that affect the sustainability of the project – hopes for future cervical cancer screening initiatives were placed on the Nicaraguan Ministry of Health, but as they were not adequately implemented for sustainability, replication of the intervention may also be difficult. The Ministry of Foreign Affairs of the Netherlands can learn from the Ixchen intervention’s failure to involve the Nicaraguan government and make a concerted effort *during the planning phase* of projects to involve relevant governments for sustainability in all intervention implementations going forward. The [Ministry of Foreign Affairs of the Netherlands website](#) [states](#) that “the Netherlands’ guiding principle is to work with local institutions, so that activities can be continued after funding has ended”.¹⁷ The hope now is for the Netherlands to follow through on this promise.
3. The third ex-post evaluation is called “[Post-mission evaluation of the integrated police training mission in Kunduz](#).” The mission in the Kunduz province of Afghanistan took place between 2011-2013; the evaluation was conducted in 2018-2019. The intervention aimed to train junior police officers, mentor middle-/higher- ranking officers, and fund projects that hoped to improve the justice system in Kunduz.⁸ Regarding sustainability, a summary of the full report noted the following:

Positive findings:

- “... there has been improvement in the professional skills, professional behaviour and knowledge of police and justice officers in the province, but these improvements cannot be attributed one-to-one to the Dutch mission between 2011 and 2013. They are related to various factors, including the overall higher level of education in Afghanistan, pressure from local NGOs and civil society activists, efforts by the Afghan government, and many international missions and programmes, including the Dutch one”.⁸
- “The Regional Police Training Centre (RTC) near Kunduz city is still functioning. Its curriculum is broadly the same as in 2011-2013, supplemented with extra weapons training. There is also a mobile training team that has been inspired by the Dutch POMLTs (Police Mentoring and Liaison Teams)”.⁸

- “AIBA (the Afghan Independent Bar Association supported by the Netherlands) is still functioning in Kunduz city. The number of registered lawyers in Kunduz rose from 21 in 2011 to 101 in 2018, including 14 women”.⁸

Negative findings:

- “... since the withdrawal of the international community in 2013, the functioning of the police and the justice system has deteriorated again, partly due to the worsening security situation”.⁸
 - “Partly because of the resurgence of the Taliban and other armed groups, Dutch focus areas such as community policing and human rights – including women’s rights – have been relegated to the background”.⁸
 - “Other factors contributing to the limited progress in Kunduz are widespread corruption, nepotism and interference by politicians and other influential players in the police, the Public Prosecution Service and the judiciary. Furthermore, there are cultural barriers that make access to justice and equal treatment in the justice system more difficult, particularly for women”.^{8*}
4. In the series of 10 evaluation publications of infrastructure projects, 19 projects were evaluated, 11 of which were evaluated ex-post.⁶ Seven of the 10 evaluation publications contained at least one project that was evaluable ex-post. For instance, in the “Drinking Water Infrastructure Projects in Ghana and Sudan” evaluation, two projects were actual ex-post evaluations while a third Ghanaian project and the two Sudanese projects were still being implemented at the time of evaluation. All the projects were funded under a program called the Development-Related Export Transactions Programme (Ontwikkelings-Relevante Export Transacties; ORET, more details of which are available [here](#)¹⁸), and the evaluations were commissioned to evaluate ORET operations from 2007-2012; the evaluations were published in 2015. If a project was completed in that 2007-2012 time period (there are a few exceptions to this, including a railway project in Bangladesh that ended in 2005), it was evaluated ex-post; if not, the project was evaluated midline or endline. Some of the important findings from projects evaluated ex-post are highlighted below.
- A) [Railway Signaling and Interlocking in Bangladesh](#) (project ended 2005) – The project designed, supplied, installed, tested, and commissioned a signaling and interlocking system of seven train stations in the Ishurdi–Jamtoil section of the Bangladesh West Zone railway.⁹ The evaluation notes financial concerns for the future as Bangladesh Railway, the partner on this project, attempts to sustain the signaling and interlocking system: “It is expected that the gap

between earnings and the resources needed for investments and maintenance of existing infrastructure will not disappear in the near future. This might have a negative effect on the future maintenance of the project equipment”.⁹ That said, the new railway system uses equipment that does not require much maintenance compared to the equipment older systems used.⁹ Furthermore, the ex-post found that the employees of Bangladesh Railway “are handling the equipment properly and efficiently, which enhances the sustainability of this equipment”.⁹ The evaluation found that maintenance sustainability was strong (even though finances are a concern) and the railway system was functioning smoothly at the time interviews were carried out in 2014.⁹

- B) [Dutch Buses for Ghana](#) (ended in 2010) – Between 2005 and 2010, 500 buses were manufactured by the Dutch and sent to Ghana.¹⁰ 475 of those buses were sent to public bus company Metro Mass Transit (MMT) while the 25 remaining buses were sent to another public bus company, STC.¹⁰ The evaluators found that 72% of buses that MMT retained ownership of were still in operation as of March 2013.¹⁰ On the other hand, “the 25 buses sent to STC wore out in just four years due to a lack of proper maintenance but ran over a million kilometers each since they were used intensively almost non-stop”.¹⁰ According to the evaluators, “the transactions have helped MMT to become a viable company”, and 43.8% of 2013 revenues still came from the Dutch-built buses.¹⁰ MMT has been able to buy its own buses without grants due to the success of the company.¹⁰ Nonetheless, the greatest ex-post concern for sustainability is that “MMT currently has few qualified staff members in technical and engineering functions”, which may be problematic for bus maintenance.¹⁰
- C) [Drinking Water Infrastructure Projects in Ghana and Sudan](#) – Two of the Ghanaian projects, called the Kwanyaku Project (fully handed over in 2013) and the Tamale Project (fully handed over in 2008), were evaluated ex-post.
- The Kwanyaku Project was split into parts I (2003-2010), II (2006-2011) and III (2008-2013).¹¹ Kwanyaku I built a water purification system and pumping stations and expanded the existing distribution network in the catchment area.¹¹ Kwanyaku II extended the network to Senya Bereku.¹¹ The phase III concurrent Kasoa Interconnection Project installed a transmission and distribution network and constructed a 500m³ water reservoir, small-scale public sanitation facilities and standpipes, to improve the water supply in catchment communities.¹¹ In total, “122 public standpipes were constructed to supply the population with piped water”.¹¹ The evaluators were concerned with the maintenance of the standpipes used to deliver water by the state-owned company in charge of the pipes, Ghana Water Company

Limited (GWCL).¹¹ The ex-post found that 42.6% of GWCL customers felt GWCL responses to “request[s] for technical assistance after [breakdowns]” were either “slow” or “very slow”.¹¹ In a similar vein, poor spare part management was also a concern for sustainability.¹¹ More encouragingly, “regarding the poverty aspect of the programme, the ORET standpipes [gave] poor people cheaper access to water”.¹¹

- The evaluation describes the Tamale Project, which was handed over to the GWCL in 2008, as follows:

“The Tamale project rehabilitated the existing water plant (20,000m³/day) near Tamale and constructed a new plant with a daily production capacity of 25,000m³ located next to the existing plant.... In addition, a new reservoir, a new transmission station and distribution pipelines were constructed.... The existing transmission pipeline supplying the corridor villages was also rehabilitated and water meters were installed”.¹¹

- Maintenance was the greatest concern to the evaluators, as the ex-post evaluators found training and maintenance assistance only lasted for six months, which did not build an adequate ‘culture of maintenance’.¹¹ Furthermore, local managers of the plant did not want to attend courses that would have taught them maintenance strategies (no year was given for when these courses were offered).¹¹ Hygiene conditions in the plant thus were lackluster (likely stemming from the lack of maintenance).¹¹

D) [One Buoy Tender and Three Aid Vessels to Indonesia](#) (completed in 2009) – The evaluation found strong results. It notes that the project supplied three vessels to Indonesia, in addition to training and technical management support, and the transfer of knowledge to a local shipyard.¹² The evaluators found good financial sustainability systems for the three vessels; the Ministry of Transportation “includes their maintenance cost in its annual budget proposal to the Ministry of Finance”.¹² In fact, the evaluators found no notable sustainability concerns in financial, maintenance, institutional, and environmental sustainability.¹²

E) [Upgrading the Disaster Response Network in the Western Province and Selected Urban Areas in Sri Lanka](#) (implemented from 2006-2011) –

The project was aimed at strengthening 18 fire brigades in 15 Sri Lankan cities.¹⁵ The project trained local fire brigades to handle complex emergencies and provided them with the necessary equipment and vehicles.¹³ The project also helped Colombo’s Municipal Council Fire Brigade set up a Special Response Unit (SRU) and a national Emergency Response Training Centre (ERTC).¹³ Numerous urgent deficiencies in the disaster response network were identified, e.g. local fire brigades were short of fire, rescue, medical equipment, and

vehicles.¹³ Coordination among local fire brigades was insufficient because of a shortage of a well-functioning organizational and physical communication infrastructure and communication equipment.¹³ Furthermore, the training necessary to receive emergency calls from civilians was insufficient.¹³

Regarding sustainability, 17 of 18 local governments were unable to fully cover the necessary costs of funding adequate fire brigades following the departure of Dutch aid when the evaluation was conducted three to four years after project end.¹³ “The shortage of finance for operation and recurrent costs... resulted in a shortage of required staff, maintenance and repairs”,¹³ and there was a lack of spare parts and servicing issues for vehicles, such as worn brake pumps.¹³

Further,

- “Most firefighters and service crews, while appreciating the contribution of the project, pointed out that fire brigades do not receive sufficient attention from local authorities.... the staff of all fire brigades visited were of the opinion that the fire brigades would be better off if they were under the aegis of a department of the Ministry of Provincial Councils and Local Governments. Such institutional embedding would also enable the staff to have a career path, the lack of which is currently a source for disappointment and demotivation”.¹³

The sustainability blame seems to rest on the shoulders of the local governments’ lack of interest to support its fire brigades, which should have been dealt with at the funding/design phase.

F) [Airport Rehabilitation in Tanzania](#) – According to the evaluation, “the project intended to fully rehabilitate the [Julius Nyerere International Airport (JNIA)] airport, aiming to meet the minimum safety requirements of the International Civil Aviation Organization (ICAO) and facilitate growth in passenger and cargo transport”.¹⁴ The intervention involved six separate ORET grant projects, five of which had ended by the time the evaluation was completed – therefore, a project that was still being completed was evaluated ‘ex-post’.¹⁴ Details for the five completed transactions are as follows: for fire and rescue systems (1999-2003), for power supply (2003-2009), for a runway upgrade (2004-2009), for airport rehabilitation supervision (2006-2010), and for a second runway upgrade (2008-2012).¹⁴ The evaluation did not go into much detail about the sustainability of these five investments (there was a particular scarcity of quantitative data related to sustainability), but the evaluators found that “the technical sustainability of the project is assured by the transfer of knowledge on how to maintain systems and the possibility of ordering spare parts”.¹⁴ The evaluators were also optimistic

about the financial sustainability of the project, stating that the rehabilitation project had “improved the airport’s capacity and has facilitated a growth of passengers and cargo transport, which raises the budget to be spent on investments and maintenance”.¹⁴

Furthermore, the evaluation states that the government has placed an emphasis on the maintenance and development of the airport, guaranteeing funding for the foreseeable future.¹⁴ These point to a thorough handover to local partners, which is an excellent finding.

G)

H) [Rehabilitating Diagnostic Services in Tanzania](#) (implemented from 1998-2006) – The evaluation described the project as follows:

- “The project aimed at contributing to the improvement of the quality of health care services by halting the deterioration of diagnostic services in Tanzania. This was realised by rehabilitation and replacement of existing, often hazardous, equipment, an extension of the coverage of the services through the introduction of new equipment, and an improvement in the quality and quantity of the necessary medical and non-medical staff through training and education on the use of the equipment”.¹⁵

According to the ex-post evaluators, though the implementation of new equipment was successful, the success was only partly sustained:

- “Financial and economic sustainability of the diagnostic services in governmental hospitals was not guaranteed. In most hospitals, the ORET ultrasound equipment was replaced by other equipment provided by other donors or hospitals own sources. Consumables like X-ray films and chemicals were provided by MoHSW [the Tanzanian Ministry of Health and Social Welfare] through the Central Medical Stores and in case supplies were insufficient, consumables were bought locally using the hospital budget. This arrangement is not sustainable due to irregularities in the supply system and flow of funds from government to the facilities”.¹⁵
- As for technical sustainability, though there was a post-implementation maintenance contract between MoHSW and Philips Medical Systems BV (the Dutch company that helped with implementation), MoSHW accumulated such a large debt – over 6 million USD – that Philips Medical Systems BV stopped providing services in 2014.¹⁵ As such, equipment has been falling into a state of disrepair.
- Similar to the fire brigade intervention in Sri Lanka, the Netherlands cannot shoulder much blame for the difficulties encountered in sustaining the outcomes of the project. That said, the assumptions made in both cases about national government capacities to sustain the equipment/interventions can be faulted.

The Netherlands Paper Citations

1. IOB. (2009, October). Evaluation policy and guidelines. Retrieved June 13, 2020, from <https://english.iob-evaluatie.nl/publications/regulations/2009/10/01/evaluation-policy-and-guidelines>
2. OECD/DAC Network on Development Evaluation. (2019, November/December). OUTLINE OF PRINCIPLES OF IMPACT EVALUATION. Retrieved July 13, 2020, from <http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf>
3. JICA. (n.d.). Ex-post Evaluation (Technical Cooperation). Retrieved October 1, 2020, from https://www.jica.go.jp/english/our_work/evaluation/tech_and_grant/project/ex_post/index.html
4. IOB. (2019, July 1). Less Pretension, More Realism: An evaluation of the Reconstruction Programme (2012 - 2015), the Strategic Partnerships in Chronic Crises Programme (2014 - 2016) and the Addressing Root Causes Tender Process. Retrieved June 8, 2020, from <https://english.iob-evaluatie.nl/publications/evaluations/2019/07/01/428---iob---an-evaluation-of-the-reconstruction-programme-2012-2015-the-strategic-partnerships-chronic-crises-programme-2014-2016-and-the-addressing-root-causes-tender-process>
5. IOB. (2012, October). Impact evaluation of the Ixchen Centre for Women cervical cancer programme in Nicaragua – Turning a right into practice . Retrieved June 13, 2020, from <https://english.iob-evaluatie.nl/publications/evaluations/2012/10/01/371---impact-evaluation-of-the-ixchen-centre-for-women-cervical-cancer-programme-in-nicaragua-2005-2009---turning-a-right-into-practice-kopie>
6. IOB. (2015, July). Country case studies – Investing in Public Infrastructure. Retrieved July 27, 2020, from <https://english.iob-evaluatie.nl/publications/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies>
7. JICA. (n.d.). Ex-post Evaluation (ODA Loan). Retrieved July 28, 2020, from https://www.jica.go.jp/english/our_work/evaluation/oda_loan/post/index.html
8. IOB. (2020, May 08). Post-mission evaluation of the integrated police training mission in Kunduz. Retrieved January 3, 2021, from <https://english.iob-evaluatie.nl/publications/reports/2019/11/19/432-kunduz>
9. Ahmed, B. N., Van der Windt, N., & Van Rijsbergen, B. (2015, February). Railway Signalling and Interlocking in Bangladesh. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Railway_signalling_and_interlocking_in_Bangladesh_201502.pdf
10. Otgaar, A., Agradi, M., Witte, J., & Genee, O. (2014, November). Dutch Buses for Ghana. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Dutch_buses_for_Ghana_201411.pdf
11. Gross, E., Senadza, B., Hodey, L. S., & Abdalla, S. (2015, February). Drinking Water Infrastructure Projects in Ghana and Sudan. Retrieved July 27, 2020, from <https://english.iob->

- [evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/University_Bayreuth_Drinking_water_infrastructure_projects_Ghana_Sudan_201502.pdf](https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/University_Bayreuth_Drinking_water_infrastructure_projects_Ghana_Sudan_201502.pdf)
12. Jasmina, H., Soetjpto, W., & Van der Windt, N. (2015, March). One Buoy Tender and three Aid Vessels to Indonesia. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Vessels_Indonesia_201503.pdf
 13. Sedere, U. M., & Van der Windt, N. (2015, February). Upgrading the Disaster Response Network in the Western Province and Selected Urban Areas in Sri Lanka. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Upgrading_disaster_response_network_Sri_Lanka_201502.pdf
 14. Kamanzi, A., George, V., Otgaar, A., & Witte, J. (2015, March). Airport Rehabilitation in Tanzania. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Airport_rehabilitation_Tanzania_201503.pdf
 15. Severens, J., Chitama, D., & Van Rijsbergen, B. (2015, April). Making Things Better: Rehabilitating Diagnostic Services in Tanzania. Retrieved July 27, 2020, from https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/sub-studies/2015/07/01/investing-in-public-infrastructure-country-case-studies/EUR_Rehabilitation_diagnostic_services_Tanzania_201504.pdf
 16. Government of the Netherlands. (2014, September 19). New fund for victims of conflict-related violence and natural disasters. Retrieved July 6, 2020, from <https://www.government.nl/latest/news/2014/09/16/new-fund-for-victims-of-conflict-related-violence-and-natural-disasters>
 17. Government of the Netherlands. (2018, May 18). Investing in Global Prospects. Retrieved June 18, 2020, from <https://www.government.nl/topics/development-cooperation/documents/policy-notes/2018/05/18/investing-in-global-prospects>
 18. Government of the Netherlands. (2015, July). IOB - Work in Progress - Evaluation of the ORET Programme: Investing in Public Infrastructure in Developing Countries. Retrieved July 28, 2020, from <https://www.government.nl/binaries/government/documents/reports/2015/07/01/iob-work-in-progress-evaluation-of-the-oret-programme-investing-in-public-infrastructure-in-developing-countries/work-in-progress--evaluation-of-the-oret-programme--investing-in-public-infrastructure-in-developing-countries.pdf>

VALUING VOICES

BETTER EX-POST EVALUATION WHITE PAPER SERIES

PART 3: Toward Building Better Evaluation Practices: A Series on the Ex- Post Evaluation Practices of Nordic Countries –Norway

By Preston Stewart, FDR Foundation/Harvard Research Intern at Valuing Voices

This paper will cover government sponsored ex-post evaluations from the Norwegian Agency for Development Cooperation (NORAD), the developmental aid department for the Norwegian Ministry of Foreign Affairs. As noted in the first paper of the series, the key findings for NORAD were as follows:

- Though NORAD ostensibly had six ex-post evaluations, only two of them were actual ex-post evaluations.
- NORAD's two ex-post evaluations were commissioned to lend specific knowledge to future projects in the areas and sectors in which the projects that were evaluated ex-post took place. The intentionality of these ex-post evaluations to guide future projects is a positive development that should be replicated in upcoming ex-post evaluations.*
- There were nine published documents calling for ex-post evaluations in the future, the most of any of the five Nordic countries researched. However, seven of those nine calls for ex-post evaluations were from one initiative, the UN-spearheaded REDD+ initiative.
- NORAD has published contradictory definitions of the term 'ex-post evaluation', suggesting that definitional issues have led to the misclassification of midline and endline studies as ex-post evaluations.

What are ex-post evaluations? [According to the Japan International Cooperation Agency \(JICA\)](#), the global leader in conducting ex-post evaluations, ex-post evaluations are generally conducted years after project completion and aim at “deriving lessons learned and recommendations to improve the project as well as to help plan and implement more effective

and efficient projects”.¹ This series of analyses of ex-post evaluations aims in part to spark a conversation about definitional accuracy within the Nordic development community.

Definitional Findings

NORAD has published two definitions of the term ‘ex-post’ that directly contradict each other. In 2009, a document called “[Evaluation of Norwegian Development Co-operation in the Fisheries Sector](#)” was published, stating that “ex-post evaluations are carried out at regular intervals after the programme or project has started”.² Under this definition, midline and endline evaluations would not be considered ex-posts. The four purported ex-posts that were not actual ex-post evaluations were mis-defined. On the other hand, another document published by NORAD in 2009, an “[Organisational Review of the Norwegian Refugee Council \(NRC\)](#)”, states that the NRC conducts “‘ex-post evaluations’ taking place some years after a project or programme has finished in order to assess impact and sustainability”³, indicating that there is an accurate definition of the term ‘ex-post’ circulating within NORAD. The agency should officially codify a version of the second definition of the term ‘ex-post’ so that future evaluations are accurately classified.

Categorization Findings

Four of the six purported ex-post evaluations found are either midline or endline evaluations that are misclassified as ex-post evaluations:

- “Ex-Post Evaluation of the Mongolia Programme”⁴ – This evaluation was published in 2006, the same year that the Mongolia (child rights) Programme ended, indicating that this evaluation was an endline evaluation rather than an ex-post evaluation. An actual ex-post evaluation would have taken place at least two years after the end of the program.
- “Organisational Development and Other Initiatives for the Deaf in Palestine: An Ex-Post Evaluation”⁵ – This evaluation was published in 2005 even though the evaluation notes that the projects were slated to end on December 30th, 2005. Because the intervention was ongoing during the evaluation, this is a midline evaluation, not an ex-post.
- “The Norway-India Partnership Initiative Phase II: Impact Evaluation of Five Interventions”⁶ – The evaluation was published in 2018 and was conducted on a series of projects ending 2021. Though not in the title of the report, the evaluation text incorrectly calls itself an ex-post even though it was a pre-final evaluation.

- “End of Project Evaluation: Education and Rural Development Phase II”⁷ – The project was slated to end in December 2011, and the evaluation report was published in the same month, indicating that this was an endline evaluation even though the executive summary calls the evaluation ex-post.

Neither of the two actual ex-post evaluations found is called an ‘ex-post evaluation’ in its title. Both are incorrectly titled ‘impact evaluation/assessment’, which is terminology that has a distinct meaning from ‘ex-post evaluation’, according to the [OECD](#).⁸ The two actual ex-post evaluations’ titles are:

- “[Impact Evaluation of Three Projects in Pokot, Kenya](#)” (the first project ended in 2002, the second ended in 2007, and the third in 2006; the evaluation report was published in 2009)⁹
- “[Impact Assessment of Norwegian Support to the Energy and Road Sectors in Pemba \(Zanzibar\)](#)” (the projects ended by 2012; the evaluation report was published in 2015)¹⁰

The titles and evaluation approach of future ex-post evaluations should not include the term ‘impact evaluation/assessment’, but should instead refer to the evaluations as ‘ex-post’. Clear, accurate categorization is important to increase accessibility and accountability.

Search Findings

Of the five Nordic databases searched, the NORAD database was the most difficult to navigate. Though helpful document previews that showed the location of the search term were available, searches often yielded numerous irrelevant results. For example, when searching the term ‘impact evaluation’, results were returned that merely contained the word ‘impact’, meaning that most results were unrelated to impact evaluations. Though improvements to the database could be made, a more useful tool than a database could be created: an online repository of ex-post evaluations that is easily accessible via a link on the IOB website, akin to [this one](#) by JICA.¹¹

Ex-Posts Found

NORAD published two ex-post evaluations in its database. This section includes a list of the ex-post evaluations found. Summaries of the findings from the ex-post evaluation reports related to the *sustainability* of results can be found in Annex 1.

1. [Impact Evaluation of Three Projects in Pokot, Kenya](#)⁹
2. [Impact Assessment of Norwegian Support to the Energy and Road Sectors in Pemba \(Zanzibar\)](#)¹⁰

Recommendations

NORAD should properly classify its reports according to the accurate definition of evaluation types. If an evaluation is actually an ex-post, it should be titled “ex-post evaluation”. If an evaluation is an impact evaluation, evaluating the effectiveness of an array of intervention activities during implementation, its title should be “impact evaluation”. Furthermore, changes should be made to NORAD publications to remove contradictory definitions of the term ‘ex-post’. **NORAD would benefit from clear, publicly available guidance highlighting this commitment to careful definitions and should create a centralized online location to access ex-post reports focused on sustainability, differentiated from other kinds of evaluations.** This way, it would be easier to learn from different evaluations so that the findings can yield clear lessons to be used for future projects.

NORAD should expand its ex-post evaluations to inform future projects. The ex-post evaluation on the electrification and road rehabilitation projects in Tanzania could be used as a model for future ex-post evaluations: there was **comparison to baseline data and a clear focus on sustainability/ex-post developments.**

The findings from the specific ex-post evaluations have led to the following recommendations:

- The projects in Pokot and the church’s difficulty to sustain outcomes from the intervention show the importance of carefully planning for exit. Though results from education, agricultural, and sanitation improvements were sustained by individuals and local committees, a more ideal situation would have been for the church to play a key role in project sustainability since they were the organization planned to take over. **NORAD should, in the future, choose its partner organizations by considering their *post-project capacities*.** The church, with its limited funds, may not have been the best choice to take over from NORAD. On the other hand, the Tanzanian government was an appropriate choice to take over road maintenance in the Pemba intervention due to the government’s role in funding infrastructure. **More broadly, projects should carefully plan exit strategies during the planning phase to increase the likelihood of smooth handovers to local authorities at the end of interventions.**
- Both the Pokot and Pemba interventions showed weaknesses in addressing women’s rights even though improving the economic and social wellbeing of women were goals of the two interventions (improving opportunities for women was a secondary goal of the Pemba

intervention).¹⁰ The Pokot evaluation found that women who were interviewed were often unaware of their human rights, and economic interventions, such as teaching women how to make handicrafts, were unsuccessful because women could not find a market for their goods.⁹ Furthermore, adult classes offered by the intervention did not even reach many women.⁹ The Pemba evaluation found that though women had experienced some economic and social gains, these changes were coming slowly.¹⁰ Although it is difficult to change traditional values around gender rights, it is possible for NORAD to make a stronger commitment to gender equality. **Material interventions that aim to improve the economic and social wellbeing of women, like offering far-reaching courses to women that would have been helpful during the Pokot intervention, should be planned before implementation begins so that sustainable, positive outcomes for women can not only be a goal for projects but a reality moving forward.**

- Finally, **ex-post evaluations should collect quantitative as well as qualitative data, beginning with the collection of baseline data on M&E-relevant indicators that can be monitored throughout projects.** The Pokot ex-post did not cite any quantitative data, so statements about project successes are purely qualitative, hence anecdotal. On the other hand, the Pemba ex-post used quantitative data comparison with baseline data, which provided strong evidence to back up claims of success. Mixed-method data collection should become the norm for NORAD's ex-post evaluations to foster transparent learning.

Endnotes for Norway

Norway's commitment to sustainable development:

[From the government of Norway](#) – “The Government considers international efforts to bring about sustainable development to be of vital importance. Sustainable development also requires efforts at national level.”

Annex 1 for Norway

- 1- The first of the two ex-posts is called “Impact Evaluations of Three Projects in Pokot, Kenya”. The three project evaluation was a combined report of the Pokot Development Programme (PDP; 1991-2002), the Pokot Integrated Programme (PIP; 2003-2007), and the Training of HIV/AIDS Community Counsellors program (2004-2006).⁹ The ex-post evaluation was

published in 2009. The goals of the PDP and the PIP (which was successor to the PDP and was a similar program) were to help reduce poverty by community and household mobilization and the use of locally available resources, to increase food security, to improve the social and economic living conditions of women, to improve individuals' health standards, and to strengthen primary and secondary education.⁹ The results were mixed.

Positive findings:

- “The educational activities, mainly construction of classrooms, seem to be the most successful and the most sustainable part of the development activities in the [PDP/PIP] area.... School committees [who now run the classes] are searching for more support from [the] Government as well as NGOs to expand their school building[s].... In the remote areas..., the communities are slowly appreciating the necessity to educate their children and their role in contributing towards education for the girls. This is happening in spite of almost 100% adult illiteracy”.⁹
- The ex-post evaluation found that the PDP/PIP's improved crop yields and animal health were sustained.⁹ On the evaluators' return visit, the soil erosion situation had improved – less erosion was observed.⁹ The evaluators saw terraced farms that had resulted in increased crop production and greater access to timber products and firewood.⁹ Agricultural lessons taught by the project continued to be used by farmers.⁹
- Participants reported that there was also a positive shift in attitudes toward sanitation.⁹ The evaluators found on their return visit that households were practicing better hygiene, such as creating greater separation between people and livestock, boiling drinking water and milk, and using raised dish racks to dry dishes after cleaning.⁹ According to the evaluation, “these activities are highly sustainable since most target groups revealed a high [and sustained] level of knowledge of dealing with hygiene and sanitation issues”.⁹ Furthermore, many areas that previously had issues with hygiene-related illnesses experienced lower rates of such diseases than before the intervention, which was a visible effect of sustained hygiene practices.⁹
- “Summing up on the sustainability of PDP/PIP, the education activities are continuing [to be] led by local committees[. The] agricultural, sanitation and health activities have made a fundamental change in the area and people continue using what they have learnt [from the intervention]”.

Negative findings:

- “The church [that handled some project takeover] has not been able to continue [its agricultural, sanitation, and health activities]”.⁹ In other words, though education, agriculture,

and sanitation outcomes have been sustained by *individuals* and *school committees*, the church has not been able to sustain its programs – in part due to limited funds, according to the evaluation.⁹ Much of the onus of sustainment has fallen onto the school committees and individuals, who have been successful at sustaining positive outcomes.⁹

- Though education, agricultural, and sanitation/hygiene activities have been sustained by individuals and local committees, sustainable HIV/AIDS training results were not achieved:⁹ “The HIV and AIDS programme did not manage to reach the goal of training pastors and counsellors at a parish level so they could continue the teaching. Due to lack of funds; it stopped before getting that far.... The initial hope for more sustainable teaching on HIV and AIDS to continue in the region without the programme is not realistic”.⁹
 - It is worth noting that initiatives intended to increase the economic and social wellbeing of women did not achieve the desired results of the project either.⁹ While the difficulties the church has had sustaining positive outcomes cannot be blamed on NORAD, there should be greater care taken to select organizations that have the capacity to sustain project outcomes and activities – the onus to sustain outcomes should not have fallen onto individuals if the project was handed over to the church. NORAD should also plan to set these organizations up for success during its planning phase; the HIV/AIDS project should have been planned so as to not run out of funding before completion.
 - It is important to mention that there was no quantitative data in this evaluation. Future ex-posts should attempt to collect baseline, endline, and ex-post mixed method – both quantitative and qualitative data.
- 2- The second ex-post evaluation is called “Impact Assessment of Norwegian Support to the Energy and Road Sectors in Pemba (Zanzibar)”.¹⁰ The objectives of the project were to provide electrification to 159 villages on the islands of Unguja and Pemba, to build a power cable connecting Pemba to mainland Tanzania (specifically to the Tanga region), and to build 44.8 km of secondary economic roads and access roads.¹⁰

Positive findings:

- On the evaluators’ return visit to Pemba two years ex-post, they found that “improvements in electricity supply and road transport have had a sizeable effect on the access to health facilities, and possibly also on the quality and affordability of health services. This has resulted in better health outcomes for Pemban women and men”.¹⁰

- Regarding the Tanga-Pemba cable, the evaluators found that the supply of electricity was “adequate and fairly reliable”, indicating that local authorities had sustained the flow of electricity between Pemba and the mainland for two years, from 2012 to 2014.¹⁰
- “Together with other similar interventions, the road rehabilitation project seems to have been successful in improving the mobility of Pembroans, in particular in some of the remote rural areas of the island”.¹⁰
- As a result of increased mobility and access to electricity, new economic activities emerged, long-standing economic activities became more efficient, and economic diversification improved.¹⁰ For example, “access to reliable electrical supply... led to both improvement of production processes in existing businesses (fishing, farming, carpentry) and to the emergence of new activities (commerce of fresh drinks and ice creams), especially when combined with increased mobility”.
 - That said, some of the new businesses struggled to find markets, especially “in cases where [the businesses] remain[ed] confined within the village”.¹⁰ As such, “in spite of its sizeable economic benefits [like economic diversification], the projects do not seem to have had a significant impact on poverty reduction until now”, indicating that the economic changes did not dramatically increase wealth three years after the end of the project.¹⁰

Negative findings:

- “While electricity seems to have contributed to higher school enrollment and attendance, the evidence is more mixed regarding the effects of road improvement. In the absence of appropriate means of transport, children have not benefitted from new roads, while some have been affected by the associated hazards. All in all, educational outcomes have gradually improved in Pemba [since the projects ended], but the specific impact of the projects might take longer to materialise in general statistics”.¹⁰
 - For context, in comparison to the baseline value taken in 2009 before the intervention, the number of students admitted in standard 1 (the grade-level for 7 year-olds) in public schools was 10899; by 2014, following the intervention, the number of students admitted in standard 1 in public schools was 14043.¹⁰
- The evaluators had concerns about three particular ex-post developments:
 - High prices for electricity – “recent tariff changes have resulted in fairly high prices of electricity for the households, which also face very heavy connection fees”.¹⁰
 - “Increased mobility on roads [increased]... accident risks”.¹⁰

- “Gender relations and the role of women in daily life have started to shift as a consequence of increased use of electricity, but these changes are still limited since many of the traditional constraints on the activity of women remain unaffected”.¹⁰

Citations for Norway

1. JICA. (n.d.). Ex-post Evaluation (Technical Cooperation). Retrieved October 1, 2020, from https://www.jica.go.jp/english/our_work/evaluation/tech_and_grant/project/ex_post/index.html
2. MRAG. (2009, January). Evaluation of Norwegian Development Co-operation in the Fisheries Sector. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/vedlegg-til-publikasjoner/evaluation-of-norwegian-development-co-operation-in-the-fisheries-sector1.pdf>
3. Bain, J., & Sørnum, V. (2009, September). Organisational Review of the Norwegian Refugee Council. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/vedlegg-til-publikasjoner/organisational-review-of-the-norwegian-refugee-council.pdf>
4. Vongsaly, C., & Beauclerk, J. (2006, September). EX-POST EVALUATION OF THE MONGOLIA PROGRAMME. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/ngo-evaluations/ex-post-evaluation-of-the-mongolia-programme-save-the-children-norway-1996-2006.pdf>
5. Kvam, M. H., & Aanestad, O. (2005, July 11). ORGANISATIONAL DEVELOPMENT AND OTHER INITIATIVES FOR THE DEAF IN PALESTINE. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/ngo-evaluations/organisational-development-and-other-initiatives-for-the-deaf-in-palestine.-an-ex-post-evaluation..pdf>
6. NORAD Evaluation Department. (2018, October). The Norway-India Partnership Initiative Phase II: Impact Evaluation of Five Interventions. Retrieved June 8, 2020, from https://NORAD.no/contentassets/a34a1f04028d4948b36cff9d9f0f5069/12.18-nipi-phase-ii_impact-evaluaiton-of-five-initiatives_main-report.pdf
7. Human Education International. (2011, December). END OF PROJECT EVALUATION "EDUCATION AND RURAL DEVELOPMENT PHASE II". Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/ngo-evaluations/end-of-project-evaluation-education-and-rural-development-phase-ii.pdf>
8. OECD/DAC Network on Development Evaluation. (2019, November/December). OUTLINE OF PRINCIPLES OF IMPACT EVALUATION. Retrieved August 7, 2020, from <http://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf>
9. Kristensen, K., & Nairesiae, E. (2009, October). IMPACT EVALUATION OF THREE PROJECTS IN POKOT, KENYA. Retrieved June 8, 2020, from <https://NORAD.no/globalassets/import-2162015-80434-am/www.NORAD.no-ny/filarkiv/ngo-evaluations/impact-evaluation-of-three-projects-in-pokot-kenya.pdf>
10. The International Law and Policy Institute. (2015, March). Impact assessment of Norwegian support to the energy and road sectors in Zanzibar. Retrieved June 13, 2020, from <https://NORAD.no/globalassets/publikasjoner/publikasjoner-2015-/NORAD-collected-reviews/impact-assessment-of-norwegian-support-to-the-energy-and-road-sectors-in-pemba.pdf>
11. JICA. (n.d.). Ex-post Evaluation (ODA Loan). Retrieved July 28, 2020, from https://www.jica.go.jp/english/our_work/evaluation/oda_loan/post/index.html

VALUING VOICES

BETTER EX-POST EVALUATION WHITE PAPER SERIES

PART 4: Toward Building Better Evaluation Practices: A Series on the Ex-Post Evaluation Practices of Nordic Countries and the Netherlands –

Finland

By Preston Stewart, FDR Foundation/Harvard Research Intern at Valuing Voices

This paper will cover government sponsored ex-post evaluations from the Ministry for Foreign Affairs (MFA) of Finland, the developmental aid department for the MFA. As noted in the first paper of the series, the key findings for the MFA of Finland were as follows:

- The MFA published five purported ex-post evaluations, four of which were actual ex-post evaluations.
- Five calls for future ex-post evaluations were found. Calls were either included in project plans or end-of-project evaluations.
- Of the five government organizations researched, the MFA of Finland was the only one to both publish ex-post evaluations and publish *only* accurate definitions of the term ‘ex-post’.
- Though the definition of the term ‘ex-post’ published by the MFA was accurate, the one misclassified ex-post evaluation, which was an endline study in actuality, indicates that this definition is not known by all evaluators.

What are ex-post evaluations? [According to the Japan International Cooperation Agency \(JICA\)](#), the global leader in conducting ex-post evaluations, ex-post evaluations are generally conducted years after project completion and aim at “deriving lessons learned and recommendations to improve the project as well as to help plan and implement more effective and efficient projects”.¹ This series of analyses of ex-post evaluations aims in part to spark a conversation about definitional accuracy within the Nordic and Dutch development communities.

Definitional Findings

The MFA of Finland's Evaluation Manual states:

Ex-post evaluation refers to an evaluation that is carried out some years after the programme has been completed. An ex-post evaluation provides evidence on the longer-term impact and sustainability of the programme. Ex-post evaluations are sometimes referred to as impact evaluations. Since impact is often included as an evaluation criterion also in other evaluations, the MFA does not use the term impact evaluation exclusively when referring to ex-post evaluation.²

This is an accurate definition – ex-post evaluations should focus on project sustainability on projects that have been fully complete (including the cutoff of aid funding) for at least two years. The MFA of Finland is credited with being the only government organization in the five Nordic countries researched to have published ex-posts and an accurate, uncontradicted *definition* of the term 'ex-post' (though *misclassification* was found). The Dutch Policy and Evaluations Department published an inaccurate definition of the term, while the Norwegian Agency for Development Cooperation published contradictory definitions of 'ex-post' (one of which was accurate; the other was inaccurate). No ex-post evaluations were published by the Swedish and Danish government development agencies (SIDA and Danida).

Categorization Findings

Of the five purported ex-post evaluations, four were actual ex-posts. Though there were no definitional issues with the MFA guidance, an endline evaluation was miscategorized as an ex-post:

- “End of Project Evaluation of Concessional Credit Scheme Project: ‘Vessel Traffic Management and Information System (VTMIS) for Coastal Surveillance in the Republic of Ghana and Automatic Identification System (AIS) for Lake Volta’”³ – This evaluation, though appropriately titled an ‘end of project evaluation’, states that it is guided by an “ex-post results framework” that was drafted earlier, during the project’s planning phase.³ However, implementation of the project was ending in 2018 as the evaluation was conducted, indicating that the results framework was intended for an endline evaluation, not an ex-post evaluation as its designation inaccurately suggests.

Other than this misclassification, which is a minor issue of equating endline evaluation with ex-post evaluation in a single publication categorization issues were not found in MFA project evaluations.

Search Findings

The Finnish database returned generally relevant results upon searches. Search terms were generally highlighted as keywords in search results. However, a more useful tool than a database could be created: an online repository of ex-post evaluations that is easily accessible via a link on the IOB website, akin to [this one](#) by JICA.^{4*}

Ex-Post Evaluations Found

The MFA of Finland published four ex-post evaluations in its database*. This section includes a list of the ex-post evaluations found. Summaries of the findings from the ex-post evaluation reports related to the *sustainability* of results can be found in Annex 1.

1. [Post-Evaluation of Strengthening of Environmental Administration at the Local Level in Nepal \(SEAM\), years 2001 – 2014](#) (ended 2014)⁵
2. [Evaluation of Rural Sector Cooperation Between Nicaragua and Finland](#) (PRODEGA program ended in 2003)⁶
3. [Final Evaluation of EIBAMAZ Programme](#) (*Proyecto de Educación Inter Cultural Bilingüe para la Amazonía* (EIBAMAZ) program ended in 2012; contains final and ex-post elements)^{7*}
4. [Final evaluation of Finland's support to the education sector in Kosovo](#) (ended in 2013; contains final and ex-post elements)^{7*}

Recommendations

The MFA of Finland must hold all of its ex-post evaluations to its definition of the term ‘ex-post’. It is certainly encouraging that this definition is followed by not only the ex-post evaluations published, but also by numerous endline evaluations that include specific reasons as to why they are endline studies as opposed to ex-post evaluations. Still, **that one endline evaluation was classified as an ex-post evaluation serves as a reminder to circulate the definition of ‘ex-post’ to evaluators so that consistent standards are established.**

Finland should conduct a greater number of ex-post evaluations and make them publicly available in an accessible online database. The “[Development Evaluation Norm of the MFA](#)”^{8*}, which spells out the MFA’s evaluation process, does not mention ex-post evaluation. The MFA should seriously consider investing in ex-post evaluation in addition to its current evaluation practices so that

lessons in sustaining positive outcomes may be derived from a greater number of projects. **Ex-post evaluations should also be clearly named as such and should be separate from final evaluations.**

This is not currently the case, as the EIBAMAZ and Kosovo education intervention ex-post evaluations were found in final evaluations.

The findings from the specific ex-post evaluations have led to the following recommendations:

- Both the SEAM and PRODEGA ex-post evaluations noted concerns related to growth beyond the original projects implemented – the SEAM evaluators were concerned that the promising local environmental administration systems implemented by the intervention would be difficult to replicate and scale up, while the PRODEGA evaluators found that the Nicaraguan dairy farmers struggled to modernize dairy practices for the competitive global market. Many aspects of the two projects were sustained in the settings the projects were implemented in, but **strategies for growth beyond the original scope of the interventions should have been planned before the beginning of implementation.** The responsibility to expand projects after implementation does not fall solely on the MFA since project administration is handed over to local authorities post-project, but planning for such growth would be a beneficial step for the MFA to consider in its projects.
- **The MFA must ensure that its projects have detailed and viable exit strategies to promote a smooth transfer of project responsibility to local authorities.** While this concern was brought up only in the SEAM evaluation, it does have general applicability. A well-planned exit strategy can position local authorities for success in sustaining the positive results of an MFA-sponsored project.

Finally, though not a recommendation made due to any particular finding, **Valuing Voices believes baseline statistics should be measured using quantitative indicators decided upon at the outset of interventions, and such indicators should be monitored throughout the intervention.** This way, ex-post evaluations can measure the same indicators on return trips to the project sites and claims of success or failure can be quantified for transparent learning. The MFA of Finland may already use such standards, but it is an important suggestion that Valuing Values believes should be made to any evaluation agency.

Annex 1 for Finland

- 1- The first of the ex-posts is called “[Post-Evaluation of Strengthening of Environmental Administration at the Local Level in Nepal \(SEAM\), years 2001 – 2014](#)”.⁵ The evaluation report was published in June 2016. There were four phases of implementation, each with a different but related objective. Phase I (2001-2008) aimed to create a replicable model for “local level environmental management” that could be used by municipalities around the country.⁵ Phase II (2008-2011) had the overarching goals of “improving the state of the environment”, guiding environmentally stable development of industry, raising awareness of pro-environment practices, and improving the use of natural resources in project areas through monitoring.⁵ The third phase, the Bridging Phase (August 2011-December 2011), was an extension of the Phase II goals.⁵ Finally, the Completion Phase (2011-2014) aimed to reduce environmental contamination through improving local environmental administration, preventing industrial and chemical pollution, implementing environmental monitoring systems, and establishing central-level support to local environmental management.⁵

Positive findings:

- The evaluators found that the project “provided sustainable implementation models for the environmental administration and management for [Village Development Committees (VDCs)], [District Development Committee (DDCs)], and municipalities”.⁵ These local administrations were still operational at the time the ex-post evaluation was conducted.
- “Based on interviews with representatives of the target groups[,] the strong engagement from the [p]roject with regard to gender and social inclusion” has made gains related to inclusion successful and sustainable.⁵
- “The Environment Development Fund (EDF) and Environment Industrial Fund (EIF) were efficiently managed to expedite an improvement of the environmental administration and management, to secure sustainability of administrative structures and models, and to encourage industries to invest in environmental technology”.⁵
- “There is a good reason to believe that the Environmental Service Center (ENSC) [will remain operational] for [at least] the next 5 years. ENSC is a non-profit distributing company, established upon the basis of public-private partnership (PPP). It aims to provide the services necessary to prevent environmental pollution and carries out scientific research to support its other services. It generates enough revenues from charging for laboratory analyses to run its own operation. Moreover, the Enviro-Nepal web-portal, developed by SEAM-N, is a valuable source of environmental information and data for the environmental community of Nepal,

including scientists, students, administrators, journalists as well as international donor agencies and investors”.⁵

- “The Sample Collection Centre (SCC) coordinated by District Drinking Water and Sanitation Coordination Committee (DDWSSC) of DDC is still fulfilling its mandated duties [as of 2016]... Monitoring was successfully piloted through a water monitoring program in the project districts. However, the continuation and success of this monitoring program depend on local authorities, stakeholders, and the public and whether they utilise the tools and procedures created by the project”.⁵
- The environmental education modules developed by SEAM continue to be disseminated throughout 14 school districts.⁵
- Industries, including the tea farming industry, continue to improve energy efficiency and use pro-environment practices due to increased public pressure and awareness within the industries themselves.⁵
- The evaluation team notes that “there is high level of relevant technical knowledge and skills available at the local level to secure that the positive outcomes of the [p]roject and benefits of the project activities continue”.⁵

Negative findings:

- One hinderance to the sustainability of results was the “slow progress of the central government ministries in adopting policy initiatives developed by SEAM-N”.⁵
- “The VDCs, DDCs, and municipalities with limited financial resources are causing a risk that well-trained and knowledgeable staff looking for opportunities to work somewhere else” will leave the country.”⁵
- “Pollution-control measures and sustainable environmental management are questionable until the law enforcement, environmental legislation and regulation are in place with appropriate fiscal instruments. Much remains to be done in this respect”.⁵
- The evaluation team found that too little effort was put into planning for scaling up/replicating the local environmental administration systems.⁵ This likely hinders the post-project growth of the SEAM administration implementations.
- The Enviro-Nepal website had a small user-base at the time the evaluation was conducted because of “inadequate promotional efforts”.⁵

- Though some results were sustained, “the [evaluation team] is of the opinion that the SEAM-N never had a clearly articulated exit strategy”.⁵ The evaluation team notes that “an explicit exit strategy should have been in the Programme Documents (PDs) from the beginning, even for the first phase, and more detailed exit plans should have been a part of the Annual Work Plans, at least for the last two years. The plans should [have included] activities to sustain results, costs, timelines and responsibilities before and after end of project”.⁵ In short, the inclusion of an exit strategy would have provided a clearer framework for increasing the sustainability of project results.
- 2- The second ex-post evaluation is called “[Evaluation of Rural Sector Cooperation Between Nicaragua and Finland](#)”. The evaluation report, published in May 2013, includes an ex-post evaluation of the Livestock Development Program (PRODEGA; 1990-2003) in Nicaragua.⁶ The project was conducted in three phases and attempted to aid the development of the Nicaraguan dairy industry. The first phase (1990-1993) established four refrigerated milk-collection centers and 90 “reference farms” that were mechanized and used up-to-date livestock-raising practices.⁶ Local roads were also improved around these farms, and the project focused on the promotion of women.⁶ The second phase (1994-1998) placed a greater emphasis on poverty reduction in the farming community and areas where the farms were located, though there was still a “continual focus on livestock” during this period.⁶ Finally, the third phase (1999-2003) saw a closer adherence to the original, phase one goal of supporting dairy farm development, consolidation, and productivity.⁶

Positive findings:

- The evaluators found that “10 years after the conclusion of the project, the survival of a part of the milk cooperatives created or supported by the project has been observed, along with the strong growth and strengthening of some of them. As a result, a large number of families have been able to improve their incomes and quality of life on a long-term basis”.⁶
- The project was “geared toward the sustainability of results: the improvement of milk commercialisation, the transfer of technology, the improvement of women’s position, the mitigation of poverty and the protection of the environment”.⁶
- “The total volume of milk collected by 6 of the 7 evaluated cooperatives has increased from around 70,000 litres in 2003 to 210,000 litres in 2013, which is equivalent to a 200% increase over the ten-year period since PRODEGA’s conclusion. For the two cooperatives that

facilitated data for 1991, growth was from 6,000 litres initially up to 50,000 litres in 2003 [733% increase from 1991], the year PRODEGA concluded, and up to 140,000 litres in 2013 [2233% increase from 1991]. All these figures clearly demonstrate the impact and sustainability of what was achieved by PRODEGA”.⁶

- “The second phase introduced an element of the Gender and Development (GAD) approach, which attempted to break traditional roles and introduce non-habitual roles for women.... This field of activities, which attempted to structurally break what was socially allowed for women, has indeed yielded results – apparently sustainable – as currently, female cattle-breeders represent between 10 and 30% of the farms according to cooperatives; this information comes from interviews at the cooperatives”.⁶ Though baseline percentages of women’s participation in cattle-breeding were not reported, the evaluation report notes that women’s participation in the occupation was almost nonexistent since ranching culture is “among the most macho cultures”.⁶ It should be noted that this success with the promotion of women was possibly the most sustainable intervention related to improving women’s situations in the PRODEGA project – other pro-women aspects of the project were less sustainable.⁶

Negative findings:

- “It can be concluded that the work with women of PRODEGA was largely [related to] the reduction of (immediate) poverty within the project, while the other components were geared towards productivity and commercialisation.... the development work carried out with the poorest of the poor is not always sustainable, but rather it tends to remain on the level of poverty alleviating assistance. This has been the case with PRODEGA, particularly when it comes to matters of underlying philosophical change of the female component (first WID and then GAD). There simply was no time to consolidate the empowerment of women at the agricultural cooperatives in favour of a commercial approach”, because only two of the 13 cooperatives started in phase II (when women’s rights were prioritized) were continued into phase III.⁶ “The notable exception [to the concerning trends seen related to women’s rights] is the work done to break barriers against the participation of women in the cooperatives and their boards of directors”.⁶
- “There are concerns about opportunities for Nicaraguan dairy products in the regional and international markets and the overall sustainability of the value chain”, concerns which stem from the low quality of the milk.⁶
- Though milk production has increased, still “only 40% of milk production [was] collected in the country at the time of evaluation”,⁶ a statistic that, in concert with the low quality of milk,

suggests the dairy industry has struggled to sustain quality standards since the end of the project.

- 3- The third ex-post evaluation is called “[Final Evaluation of EIBAMAZ Programme](#)”. The *Proyecto de Educación Inter Cultural Bilingüe para la Amazonía* (EIBAMAZ) program ran from 2004-2012, and evaluation was completed in 2015.⁷ This intervention aimed to develop the capacity for the education of children from indigenous ethnic groups in the Amazonian region.⁷ The project included initiatives to create indigenous language learning materials, train teachers, and develop relationships with communities.⁷ Regarding sustainability, the following findings were discussed in a synthesis report of Finnish education project evaluations:
 - “According to interviews conducted as part of this evaluation, there are many new educators formed with EIBAMAZ who are strong advocates of their culture. There are rural networks supporting indigenous teachers, something that didn’t exist before EIBAMAZ. These individuals are now called on regularly by the Ministry of Education and by other projects to provide their expertise on IBE. In Peru in particular, materials are produced by indigenous centers on a regular basis and are available in schools, even if teachers need more support on how to use them. This is an important development... the strength of indigenous organizations is an enabling condition that will be a long-term determinant of sustainability”.⁷
 - “An important lesson for planning and implementing funding initiatives with a promise of sustainability can be obtained from the examples of Ecuador and Peru, whose paths went in opposite directions. EIBAMAZ started in Ecuador with the strong support of the government, but its achievements are now isolated, fragmented, and at risk. In Peru on the contrary, EIBAMAZ began with little official support, but has become a multidimensional effort that is likely to be sustained as a state-supported initiative. The lesson is to consider the different types of sustainability involved in a project that targets the rights of neglected social groups. Political sustainability had a good foundation in Bolivia, but it proved to be fragile in Ecuador. Technical sustainability is a growing force in Peru. Sociocultural sustainability was important for all three countries”.^{7*}
- 4- The fourth evaluation is called “[Final evaluation of Finland’s support to the education sector in Kosovo](#)”. This evaluation includes ex-post evaluation of the 2000-2003 Finnish Support to the Education Sector in Kosovo (FSDEK) and the 2004-2008 FSDEK II projects intended to

train teachers in special needs education (FSDEK) and add inclusivity training to teacher professional development programs (FSDEK II).⁷ Regarding sustainability, the following findings were discussed in a synthesis report of Finnish education project evaluations:

- “Respondents consistently state that Inclusive Education is included as a priority in the core education strategies and laws largely due to Finnish support. The FSDEK project supported development of in-country expertise and there is an operational Special Needs Unit (SNU) in the MEST [(the Kosovo Ministry of Education, Science, and Technology)]. The majority of the teachers trained by the Finnish projects work in the education sector. However, financial sustainability and the sustainability of capacity development efforts remain a concern. After Finnish support ended, the Professional Development Programmes were not continued, and Reading Clubs, which seemed to have an impact on reading skills, were discontinued. However, negotiations between with the MEST and the Faculty of Education of the University of Pristina are underway to review the Professional Development Programme with the goal that it be delivered by the newly established In-Service Training Centre”.^{7*}

Endnotes for Finland

- The claim that the definition of the term ‘ex-post’ may not be widely shared among evaluators is supported by the “Development Evaluation Norm of the MFA” document, which makes no mention of ex-post evaluation in the MFA’s development norms. This document, as well as the Evaluation Manual, was generously provided by the evaluations team of the MFA of Finland following communication between Valuing Voices and the team.
- The evaluation team of the MFA of Finland provided two more evaluations that include ex-post analysis. While the synthesis report is cited above and in Annex 1, the full-length reports can be found [here for EIBAMAZ](#) and [here for the Kosovo intervention](#).
- The evaluation team provided a link to a repository of evaluations, which can be found [here](#). However, the repository does not stratify evaluations temporally (i.e. mid-term, final, ex-post).

Finland's commitment to sustainable development:

[From the MFA of Finland](#) – “The 2030 Agenda [for Sustainable Development] includes 17 different goals that are related to sustainable development. These goals are to be met by 2030. Finland is committed to reaching them both at home and in its international cooperation. Finland's development cooperation supports this commitment.”

The evaluation team sent along two management responses (in Finnish) to ex-post evaluations, indicating that learning is occurring from such evaluations. However, a management response addressing the education efforts detailed in the “Inclusive Education in Finland's Development Cooperation in 2004-2013: Synthesis Report” document only mentioned sustainability once; the response was not focused particularly on the sustainability of the projects' benefits.

Special thanks to the evaluations team from the MFA of Finland, specifically Ms. Anu Saxen and Ms. Sanna Pulkkinen, for providing their input on this paper. The team provided the author with clarification on various points in the paper, as well as documents relevant to the analysis of this paper (including the repository of general evaluations). The paper has been edited to reflect their suggestions.

Citations for Finland

1. JICA. (n.d.). Ex-post Evaluation (Technical Cooperation). Retrieved October 1, 2020, from https://www.jica.go.jp/english/our_work/evaluation/tech_and_grant/project/ex_post/index.html
2. MFA of Finland. (2020, June 16). Evaluation Manual. Retrieved January 4, 2021, from <https://eoppiva.zapter.io/evaluationmanual2018>
3. Ministry for Foreign Affairs of Finland. (2018, August). End of Project Evaluation of Concessional Credit Scheme Project: “Vessel Traffic Management and Information System (VTMIS) for Coastal Surveillance in the Republic of Ghana and Automatic Identification System (AIS) for Lake Volta.” Retrieved June 9, 2020, from https://um.fi/documents/384998/0/FCGInt_FinalReport_PIF_Ghana_17082018.pdf/a66dcef1-6ccf-2d17-325b-9c4d4b6c1f5e#search=%22ex-post%22
4. JICA. (n.d.). Ex-post Evaluation (ODA Loan). Retrieved July 28, 2020, from https://www.jica.go.jp/english/our_work/evaluation/oda_loan/post/index.html
5. Axberg, G. N. et al (2016, June 30). Post-Evaluation of Strengthening of Environmental Administration at the Local Level in Nepal (SEAM), Years 2001 – 2014. Retrieved June 9, 2020, from https://um.fi/documents/384998/0/Nepal_SEAM_Post_evaluation%2BTOR_2016.pdf/03bc624c-2241-7fd0-2c03-0d9b3c68feda#search=%22ex+post%22
6. Seppänen, M., Eriksson, L., Aguilar, B., Boman, K., & Pijnenburg, T. (2013, May). Evaluation of Rural Sector Cooperation between Nicaragua and Finland. Retrieved June 9, 2020, from https://um.fi/documents/384998/385866/prodega_final_report/b4812a1d-5bc0-b467-3634-4fd13763cb8e?t=1528278516971
7. Nielsen, H., Prouty, R., & Bennett, A. (2015). Inclusive Education in Finland's Development Cooperation in 2004-2013: Synthesis Report. https://um.fi/documents/384998/385866/synthesis_report_inclusive_education_in_finland_s_development_cooperation_in_2004_2013
8. MFA of Finland. (2015). Development Evaluation Norm of the MFA. Retrieved January 4, 2021, from https://um.fi/documents/35732/48132/development_evaluation_norm_2015/bd2579d5-ccb3-bb0f-540d-1db0513a3a76?t=1525690790183