Evaluation Guide: Evaluating the impact of co-design for Nature-Based Solutions

March 2021
Social Finance for CLEVER Cities

Contents

1. Introduction

An evaluation framework for NBS co-design

Co-design is a key principle in CLEVER Cities, and in many Nature-based Solutions (NBS) and urban regeneration projects. **Co-design** is the involvement of stakeholders in designing or rethinking an output, through direct collaboration with the design team[1]. There is wide interest in the potential added value of this approach. However, there is very limited evidence of the impact and value for money of co-design, and this is particularly true for NBS where co-design is still an emerging practice.[1]

One of the main reasons for this limited evidence base is the difficulty of such an evaluation given the wide-ranging activities of co-design and the fact they are often part of a wider series of processes around planning. We have therefore designed a framework approach to support CLEVER cities, and other NBS initiators, by suggesting an approach to evaluation.

Why evaluate NBS co-design?

Evaluating the impact of NBS co-design can provide three key benefits:

1. **Understand:** Build an evidence base to demonstrate the impact of NBS co-design, so its benefits are understood and trusted by stakeholders
2. **Learn:** Focus stakeholders on key outcomes, to learn what works and what doesn't and improve performance on "impact"
3. **Invest:** Demonstrate the value-for-money case to invest in co-design for NBS

Who is it for?

This guide is intended for use by project teams who are planning and evaluating NBS. It may be particularly relevant to teams responsible for community engagement and co-design.

What does it cover?

- **"Co-design"**: This framework focuses on co-design. Co-design is the first element of "co-creation" – the latter also encompasses "co-monitoring" and "co-evaluation" as following steps. However, the outcomes associated with these other phases are likely to be relatively similar.
- "**Community":** Co-design involves the participation of many stakeholders in the design. This framework aims to measure the participation of the "community" – including both citizens and representatives of community organisations.
- **Results chain from "inputs" to "impact"**: Please note that, although long term outcomes (or "sustainable impact") are part of the Theory of Change, we do not provide guidance on how to measure long term outcomes. Our rationale for excluding the measurement of long term outcomes was a recognition that most NBS initiators are not ready for or do not have the resources for that level of complexity in impact measurement of co-design.

[1] In two literature reviews which evaluated 5000+ articles on co-production each, one found just 8 studies evidencing impact on the wellbeing of communities (What Works Centre) and the other found just 120 studies which evidenced any type of impact (Voorberg). What Works Centre for Wellbeing, "Community wellbeing impacts of co-production in local decision-making," 2017; Voorberg et. al, "A Systematic Review of Co-Creation and Co-Production: Embarking on the Social Innovation Journey," 2014.

However, we encourage teams to measure long term outcomes if they are already comfortable with previous steps.

Evaluating co-design in 3 steps

This report (referred to as the "Evaluation Guide") provides guidance on how to evaluate the impact of community co-design for NBS, in 3 steps:

1. Understanding what we want to achieve with co-design (section 2): this describes the "Theory of Change" of NBS co-design, including inputs, activities, outputs, outcomes and impact.
2. Measuring NBS co-design (section 3)
3. Valuing the NBS co-design (section 4) – optional, for users that may be interested in a cost-benefit-analysis based on monetisation of the impact.

This **Evaluation Guide** is intended to be read and used alongside the corresponding **Evaluation Framework** document, which:

- Provides guidance on how to measure Inputs, (quality of) Activities, Outputs and Outcomes
- Proposes a cost-benefit approach to estimating Net Added Value (Inputs, Impact) of the co-design

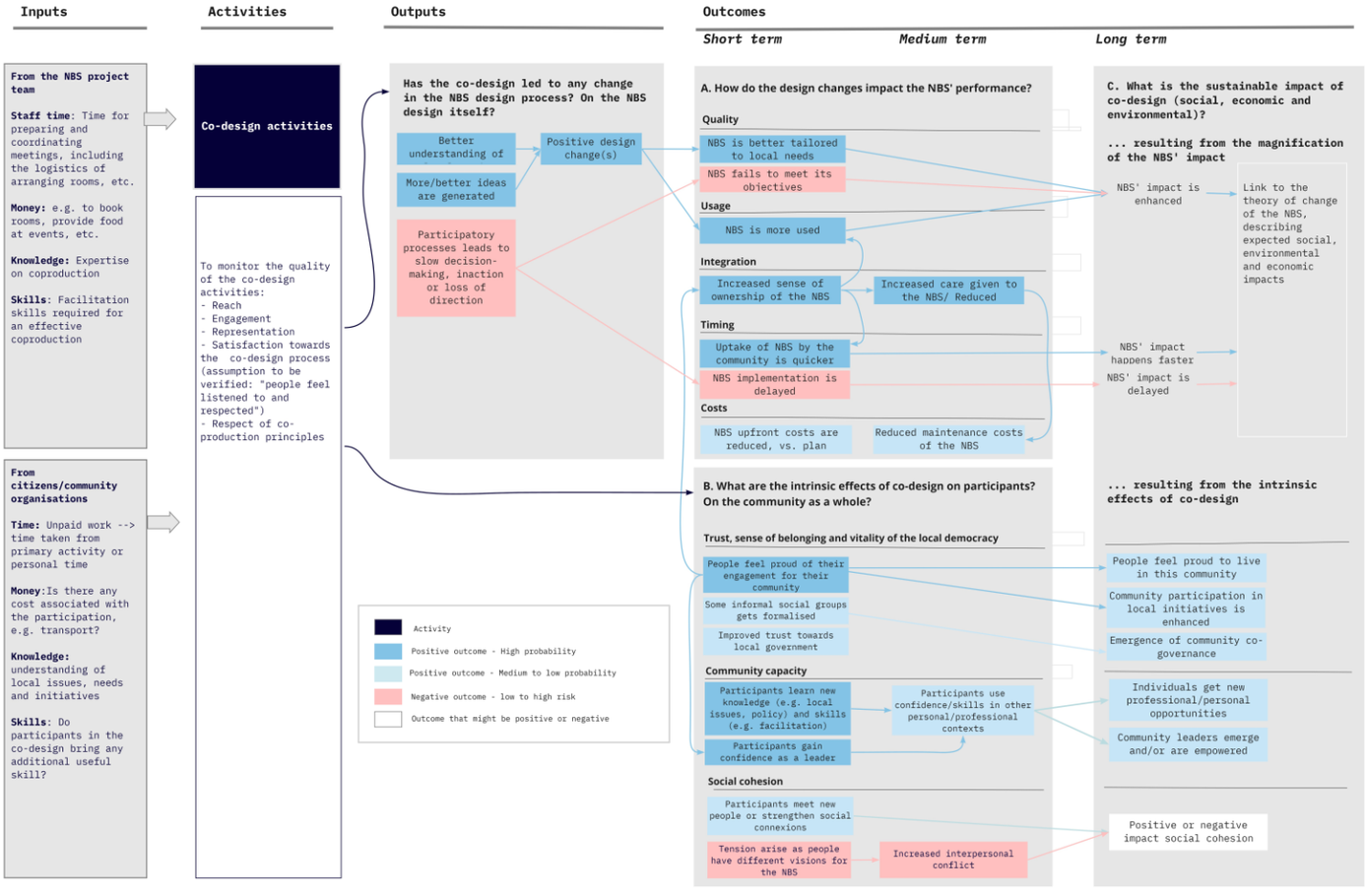
2. Understanding what we want to achieve with co-design

Before evaluating co-design, we need to understand first what we want to achieve with it. What are we expecting from it? What are the direct outputs, and the sustainable impact that we want to see?

Each NBS will have different objectives, and we invite you to define your own "Theory of Change" of co-design. However, not everything is unique to each project: there are similarities in the expected outputs and outcomes of NBS co-design.

Our framework is based on the following "generic" Theory of Change for NBS co-design (see figure 1). It describes the positive outputs and outcomes that can be expected from NBS co-design as well as the risks of negative outcomes, which are to be mitigated and monitored.

Figure 1 – Theory of Change of NBS Co-design



3. Measuring NBS co-design

Why it can be challenging

Co-design measurement poses the following challenges[2]:

- Co-design can include a wide range of activities
- Co-design is part of a wider process rather than being a distinct programme or intervention. This can make it difficult to directly link and attribute the co-design with the outcomes
- Co-design may not have pre-defined outcomes
- Different participants in co-design may engage in different ways or have different experiences

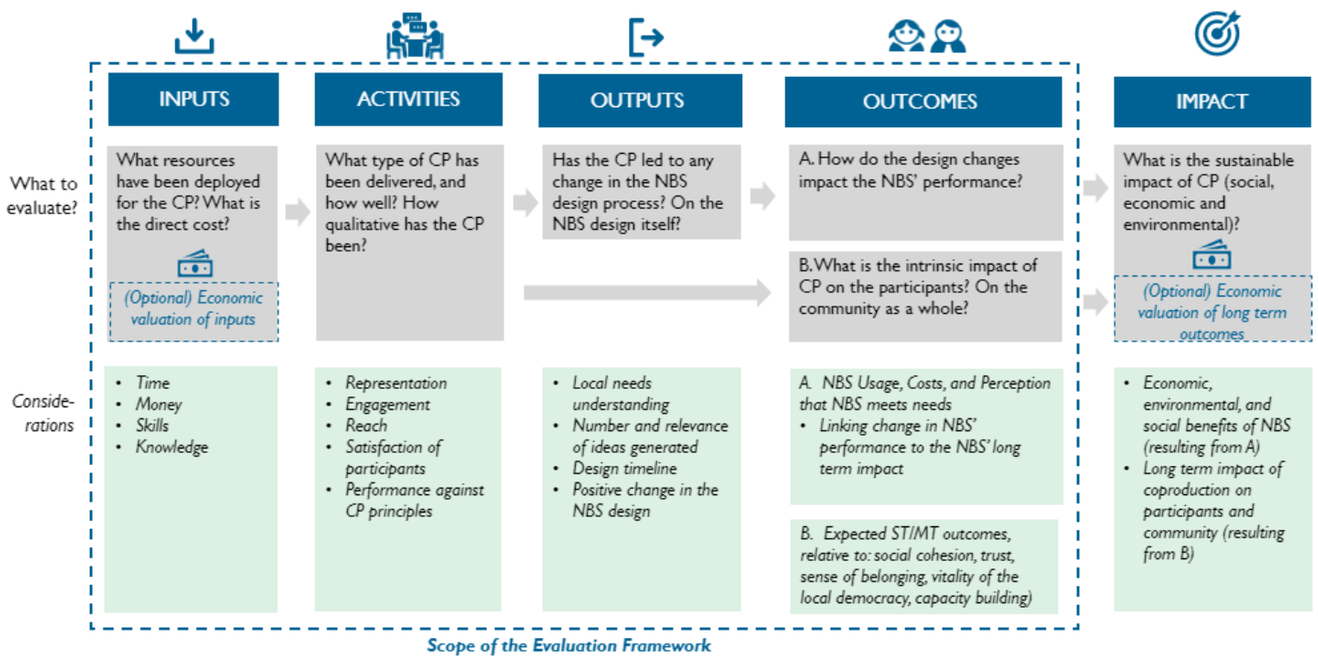
Other difficulties specific to co-designing for NBS include:

- The large number and variety of stakeholders (22 stakeholders on average, based on a study of 16 NBS[3]) involved in the co-design make this collective intervention particularly complex and unpredictable. Many variables interact[4].
- This complexity is increased by the timeline: the co-creation process of NBS often takes years - during which other changes happen in the community (other interventions, external factors such as political elections, economic crisis, etc.). Isolating what happens as a result of the co-design is challenging.
- NBS are intrinsically unique, depending on the characteristics of each territory and its community. No two NBS are alike. Hence, evaluation techniques based on the comparison between two groups (one that received an intervention, the "target group", and another group that did not, called the "control" or "comparison" group) can be difficult to implement. (See "Attribution").

Overview of our approach

Our approach to co-design measurement focuses on four areas: Inputs, Activities, Outputs and Outcomes. The corresponding Evaluation Framework offers a detailed framework to measure each of these areas.

Figure 2 – Measurement framework for NBS Co-design



The below table outlines the rationale for measurement within each of the four areas, and notes the section of the Evaluation Framework which corresponds to each. The Evaluation framework includes a range of suggested

[2] Challenges from "Implementing and evaluating co-design", NPC.

[3] Zingraff-Hamed, A.; Hüesker, F.; Lupp, G.; Begg, C.; Huang, J.; Oen, A.; Vojinovic, Z.; Kuhlicke, C.; Pauleit, S. Stakeholder Mapping to Co-Create Nature-Based Solutions: Who Is on Board? *Sustainability* **2020**, *12*, 8625.See https://www.mdpi.com/2071-1050/12/20/8625

[4] For more information, see "Evaluating the Effects of Co-Production Initiatives in Public Service Organizations", Brix et al., 2017.

measurement options (including the main inputs / activities / outputs / outcomes), key indicators, suggested approaches to data-collection and possible paths for attribution.

Section	Why measure it?	Where is this in Evaluation Framework?
Inputs	To understand what resources and other inputs the co-design requires – this supports a net value calculation of the NBS	Tab **Inputs**
Activities	To understand which elements of co-design were used, and how well the co-design was implemented	Tab **Activities**
Outputs	Measuring the design impact generated, to attribute outcomes of the resulting NBS to co-design	Tab **Outputs**
Outcomes	Key component to understand when evaluating impact	Tab **Outcomes (A)** - impact of design changes (resulting from co-design) Tab **Outcomes (B)** - intrinsic impact of co-design

How to use the Evaluation Framework?

You may not want to measure everything that is suggested in the Framework – some indicators may not be relevant to your own NBS co-design, or you may not have the resources to do so. We do encourage you to make it your own, to select what it is relevant or possible for you, and to adapt or complement it depending on your needs and realities. The Framework is intended to be used as a guide, and in no way aims to be prescriptive.

Our tip: Start small! Impact measurement can be very insightful when done properly, but it is important not to fall into the trap of trying to measure everything, even more so when resources are scarce.

When to use it?

The evaluation of co-design should be planned as soon as possible, ideally before starting any co-design activity. Hence, we invite you to look at this Framework, and define your own approach to evaluation early on in the project, whenever possible.

We suggest to collect data at different stages of the co-design process. The timing and frequency of data collection is specific to each indicator, and is specified in the Framework. For some indicators, the data collection is “ongoing” (e.g. capturing the number of participants per co-design event).

Attribution in the context of co-design measurement

What is attribution and why is it important?

It is essential that impact evaluation measures not only the changes that have occurred to a population receiving the studied intervention, but also seeks to understand the role of that particular intervention in producing these changes. This process is known as "Causal attribution". It aims to answer the question: "To what extent are observed results due to the activities of my programme rather than other factors?". Attribution is one of the most challenging aspects of impact measurement. Unlike natural sciences where researchers can work in labs to isolate different factors, in the NBS context we operate in the real world where our target population is influenced by many factors in their environment. This is even more true in complex programmes such as NBS where a large amount and a diversity of stakeholders get involved, and where the intervention may take place over a couple of years.

Our methodology

To build this framework, we have analysed all attribution methodologies described in Strategies for Causal Attribution UNICEF guide[5], in light of NBS co-design. In the “Attribution” tab of the Evaluation Framework, we have listed all methodologies we have assessed as having a “high” to “medium” relevance to our context. We explain why they are potentially relevant, and how can the approach be used in practice in our context of NBS co-design.

For each outcome, we have listed all attribution methodologies that may be used, giving you the flexibility to select the approach that is the most suitable to your context. In some occasions, we have added comments to help you think about the attribution for specific outcomes.

Three categories of attribution strategies

These methodologies are classified into three broad categories:

5 “Overview: Strategies for Causal Attribution: Methodological Briefs”, Patricia Rogers, Impact Evaluation No. 6, 2014.

1) **Estimating the counterfactual**. It aims to answer the question: "What would have happened otherwise?". This methodology is the most scientifically robust. In our context, the most scientific “experimental designs”, based on Randomised Control Trials, are not possible and relevant, and only “semi-experimental designs” are relevant.

   This approach is only possible if a "control" group can be formed, i.e. if these two conditions are met:
   - a similar type of NBS (with similar goals) is being designed in the same community, or in another community with comparable characteristics;
   - This second NBS is not being co-designed.
   Due to the unique nature of each NBS, we recognise than establishing a control group is unlikely to be possible.

In cases where a counterfactual is not possible, the two following approaches – which can be qualified as "contribution approaches" as opposed to “attribution approaches”, can help the evaluation team come to reasonably robust conclusions about the contribution being made by programmes to observed results.

2) **Assessing the consistency of evidence** for the causal relationships made explicit in the Theory of Change. This encompasses various complementary approaches, including: assessing the achievement of intermediate outcomes (and assuming that the long term impact is a result of our intervention if intermediate outcomes occur), asking people who have experienced changes to identify the causes of these changes, checking results against expert predictions, checking consistency with existing literature, and comparing with other cases studies.

3) **Ruling out alternative explanations**, through a logical, evidence-based process. If you can identify one or two factors (e.g. another intervention, or a change in political leadership) that may have had a significant contribution to the changes that you are expected or that you have measured, testing with people the extent to which they have been influenced by these alternatives factors can be a good idea.

> ***“A result of the contribution analysis does not represent the ultimate truth. It has to be regarded as a sufficient conclusion concerning the extent – and the reason(s) why – the intervention has contributed with a given outcome”*** *Evaluating the Effects of Co-Production Initiatives in Public Service* Organizations[3]

Data collection methods: how to choose between a survey, interviews and focus groups?

This framework covers mostly qualitative information, and it can be difficult to decide which method is best to capture it. For example, to measure the extent participants in the co-design “feel proud of their engagement”, should you use a survey or carry out interviews? There is not a simple answer to that. There are different ways – equally appropriate - of collecting this information and this framework is not prescriptive. Although surveys are more common, 1-to-1 interviews might be more valuable, for example if you're looking to understand in depth people's thinking or feelings. You will be best placed to decide which method is most appropriate to your context, but it is good to keep in mind the pros and cons of each data collection method, as well as good practices adapted to our context.

Type	Pros	Cons	Our Advice
Survey*	- Quick (for respondents and evaluator) - Can target a large number of people (whole population, or a representative sample) - If anonymous: lower risk of bias in answers	- Can be considered “formal” or associated with “paperwork”. It is not as gratifying as being listened to for co-design participants - Not appropriate to understand the “why” behind answers	- Keep it short - Keep it digital where possible and accesible
1-to-1 Interviews*	- Gratifying to co-design participants to feel listened to; - Helpful to understand the “why”, or to discuss sensitive / emotional subjects - Enables capture of unexpected additional information such as unexpected impact or blockage	- Time-consuming, for interviewees and for evaluator - Higher risk of bias than a survey due to the “please the interviewer” effect (mitigating bias requires the interviewer to have skills in user research)	- Must be carried out by a trusted person, able to build sense of trust and neutrality; this could be a community member if appropriate. - Start small (e.g. 5-8 interviews). Sample doesn’t need to be large, unless you need a high level of evidence. A second phase of interviews can be used if needed.

			- Important to access the "harder to reach" population - Can be done in person, over the phone or via a video conference
Focus groups*	- Time-efficient for the evaluator - In some contexts, people feel more at-ease when joined by their peers - Good to gather subjective perspectives - Best way to exchange viewpoints and discuss: group dynamic can be good to think about new ideas, discuss blockages, etc.	- Power dynamic risk, some people may not express themselves - Can be inappropriate to talk about sensitive subjects - Logistics/costs associated with method such as room or drink/food	- The facilitator needs to create an environment that encourages participants to express personal opinions and feelings

**For more information about these qualitative data collection methods in the context of evaluation, see CDC Evaluation Briefs, respectively about surveys (called "questionnaires")", interviews" and "focus groups".*

Sample

In our framework, we suggest to collect data from 3 stakeholder groups:

- "Active" co-design participants (threshold of frequency to be qualified as "active" is to be defined)
- The community as a whole
- NBS project team

The number of people to be interviewed or surveyed will depend on: 1. Your resources for this evaluation, 2. The number of people in each of these groups; 3. The level / stregnth of evidence you need.

The approach to sampling will be different depending on the stakeholder group. If interviewing every member of the community is probably not the right approach (and you should instead select a sample), interviewing all "active" co-design participants, or all NBS project team members may be possible and desirable.

Stakeholder group	Survey total population or sample?
"Active" co-design participants	Total population or sample
Community as a whole	Sample
NBS project team	Total population or sample

Most of the data you will collect is qualitative. Building a "statistically representative sample" is not very relevant for qualitative research, which instead talks about getting to a "saturation point". This occurs when repeated interviews result in the same themes and findings, or adds little to the existing information[6]. Ten interviews may be enough for you to cover most of your findings. Whilst this approach does not aim to be fully statistically "representative" of the population, it is important that you avoid bias in selecting participants in qualitative research, and that you ensure a good mix of your target population in your sample[7].

4. Valuing the NBS co-design

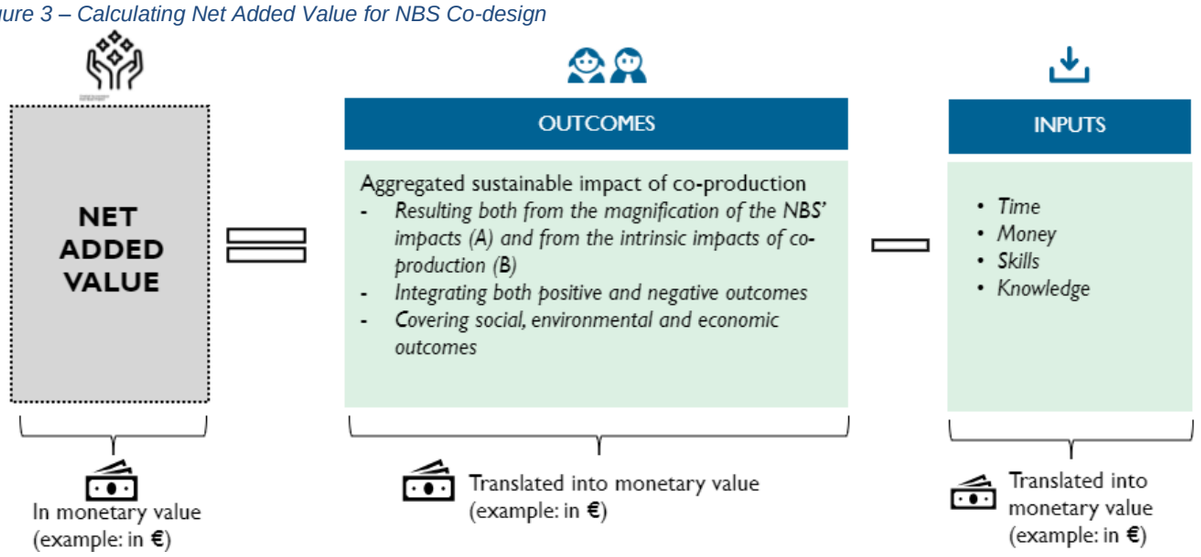
Performing a calculation of the net added value of NBS co-design can help stakeholders consider the costs and benefits of NBS co-design, and ultimately strengthen their case for using co-design in NBS projects.

Net added value analysis for NBS co-design

Our approach to estimating the net value of coproduction relies on economic valuation: each "input" and each "long term outcome" measured in the Evaluation Framework is given an estimated monetary value. The net added value is calculated by comparing the value of "long term outcomes" with the value of "inputs"; this approach is consistent with cost-benefit analysis.

[6] https://www.thinknpc.org/themes/measure-and-manage-impact/impact-measurement-evaluation-and-data/3-sampling/
[7] For more information about sampling in impact evaluation, see this article from NPC, or this article about Qualitative Impact Assessment Protocol from Better Evaluation.

Figure 3 – Calculating Net Added Value for NBS Co-design



The Evaluation Framework provides the following guidance to assist with the net added value calculation:

Topic	Description	Where is this in Evaluation Framework?
Identifying long-term impact / outcomes	The valuation will be perfomed on long-term outcomes (Impact in Figure 2). The measurement framework focuses on what is possible to measure (may be short/medium term outcomes), so we will need to identify long-term impact to perform the value calculation. This tab shows an example of how to go from the outcomes in the measurement framework to some of the longer term outcomes valued in tab "Reference – Outcome Values"	Tab **Outcomes to Impact**
Value of outcomes	This tab shows sample long term outcomes of 1. co-creation and 2. NBS. These outcomes are organised by area (economic, environmental or social). Sample monetary values are provided for each. The source of these values is related to cost-benefit and social return on investment (SROI) analyses for communities and natural capital. These values are not comprehensive - they are intended to help the reader think through the possible ways an outcome can be translated into a monetary value	Tab **Reference – Outcome Values**
Example Valuation	Example valuation, to use as a reference point and outline for a project's own co-design impact valuation	Tab **Value-Add Calculation Example**

5. Conclusion

In conclusion, evaluating co-design for NBS can be useful to build an evidence base, focus stakeholders on outcomes and "impact" and demonstrate the value-for-money case to invest in co-design for NBS. As demonstrated above, measurement can be complex. We hope that this guide has helped you consider a feasible way to evaluate the co-design for your NBS. The next step is to read through the accompanying Evaluation Framework. As discussed above, this guide corresponds to the framework, which provides guidance on how to measure each of: Inputs, Activities, Outputs, Outcomes and Net Added Value.