Funded by the Horizon 2020
Research and Innovation Programme
of the European Union

Project Number:	763784
Project acronym:	**Designscapes**
Project title:	**Building Capacity for Design enabled Innovation in Urban Environments**
Starting date:	June 1st, 2017
Duration in months:	51
Type of action:	CSA (Coordination and Support Action)
Call identifier:	H2020-SC6-CO-CREATION-2016-3
Topic:	CO-CREATION-02-2016 User-driven innovation: value creation through design-enabled innovation

Deliverable D2.3 (Part B)
Common impact evaluation methodology

Work package:	WP2
Type of Deliverable:	Report
Due date of Deliverable:	30.9.2021
Actual submission date:	31 October 2021
Dissemination level:	√ **PU (Public)** CO (Confidential)

Statement of Originality
This Deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

www.designscapes.eu

Editor's note

The current deliverable is the final deliverable of WP5. This deliverable develops the DESIGNSCAPES capacity building program and a set of training activities aimed at diffusing capacity for Design Enabled Innovation across Europe. These include the final set of Training Modules for Urban Design Enabled Innovation and 'Train the Trainer' activities.

Abstract
This document is the second part (Part B) of the final Deliverable D2.3 of Designscapes WP2 Designscapes Framework. It presents a common impact evaluation methodology which can be applied to assess design enabled innovations at project, programme and policy levels. The document first outlines a common evaluation framework, and then discusses three specific application areas (user benefits and business impacts, the relationship between design and innovation, value creating networks and their contribution to efficiency and competition).

Keywords
Common impact evaluation methodology, DEI

Copyright

Disclaimer

All information included in this Deliverable is subject to change without notice. To get the latest edition of it, the reader is advised to contact the following email address:

info@designscapes.eu

The Members of the DESIGNSCAPES Consortium make no warranty of any kind with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose.
The Members of the DESIGNSCAPES Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Authors Details

Name and Surname	Initials	E-mail address	Role (*)	Partner
Joe Cullen	JC	j.cullen@tavinstitute.org	LE	TIHR
Anna Sophie Hahne	ASH	a.hahne@tavinstitute.org	C	TIHR
David Drabble	DD	d.drabble@tavinstitute.org	C	TIHR
Kerstin Junge	KJ	k.junge@tavinstitute.org	C	TIHR

(*) LE = Lead Editor, E=Editor, C=Contributor, R=Reviewer, PC= Project Coordinator
Note: all partners read and commented the text

Document History

Version no.	Date	Author's initials	Contents and/or changes made
1	1.9.2021	JC	Impact evaluation methodology
2	29.9.2021	ASH	Regression modelling text
3	30.9.2021	DD	Systems mapping text
4	28.10.2021	KJ	Contribution analysis examples

Table of Contents

List of Figures and Tables

List of Tables

Designscapes Common Impact Evaluation Methodology

About this Document

This document sets out a methodology and provides a set of tools, accompanied by illustrative examples of their use, to support the evaluation of design-enabled innovation (DEI). The evaluation approach proposed is shaped by the provisions of the Horizon 2020 SC6-CO-CREATION-2016-2017 Call Text - Co-Creation 02-2016 ('Co-Creation for Growth and Inclusion - User-driven innovation: value creation through design-enabled innovation') [1].

Embedded in this Call Text is a set of assumptions about the connection between design-enabled innovation and competitiveness. This set of assumptions can be thought of as an over-arching (albeit basic) 'Theory of Change' for the H2020 Co-Creation 02-2016 programme. Theory of change tells the programme 'story' – from the 'presenting problem' it addresses through to the change it hopes to make on that problem at the end of the programme and beyond (i.e. the programme's expected 'impact'). It incorporates a theory of how and why the programme might cause an effect. The programme 'theory' behind the Call is shown in Figure 1.

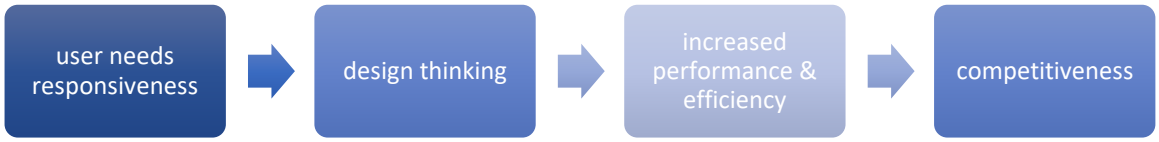


Figure 1: Co-Creation Theory of Change

As Figure 1 shows, the co-creation Call speaks to the need for Europe to become more competitive – by identifying and capitalising on untapped sources of growth and employment - in order to renew the legitimacy of public policy-making, especially through greater citizens' involvement - and to deliver better public services for all. This requires effort to encourage creativity and collaboration between various societal actors through 'co-creation' – focusing in particular on co-creation for growth and inclusion.

The path to this desired outcome of increased competitiveness starts with being responsive to user needs. User involvement is seen as central to the process of innovation. User-centred design thinking, and the use of design-thinking methods

1 https://ec.europa.eu/research/participants/data/ref/h2020/wp/2016_2017/main/h2020-wp1617-societies_en.pdf, p.11

and tools, encourages user involvement, and is expected to lead to service and business model innovation, which in turn promotes increased performance and efficiency in public, private and third sectors, ultimately increasing their competitiveness.

An evaluation methodology aimed at assessing the contribution of design thinking (or more precisely design-enabled innovation) to competitiveness therefore needs itself to speak to this over-arching 'theory of change'.

Addressing the Call Objectives

The Co-Creation 02-2016 Call Text, cited above, sets out a number of objectives, and expectations on impacts, that are relevant for the development of an evaluation methodology for design-enabled innovation. With regard to ***scope***, the Call envisages a methodology that, drawing on the experience of the projects funded by the Call, will achieve the following objectives:

a. gather data and metrics concerning the impact of design-related policies and programmes in terms of user benefit and business impact
b. develop a transferable methodology to evaluate the effectiveness of design in the innovation process
c. develop a methodology on how actors in different sectors can better connect with design-enabled innovation to increase efficiency and competitiveness in their respective sectors as well as a common impact evaluation methodology and respective indicators, to be applied across sectors and scalable to organisational, regional, national and European level.

With regard to ***expected impacts***, the Call anticipates:

- creation of data, data sets and metrics to evaluate impact regardless of sector
- results, e.g. the impact evaluation methodology, that should be applicable and give clear guidance particularly to SMEs and public administrations
- producing assessment and simulation tools to verify results in the increase of efficiency and/or effectiveness through the use of design.

These objectives and expected impacts, as expressed in the Call Text, reflect a degree of ambiguity and overlap. For example, the Call Text appears to call for two separate evaluation methodologies – one to evaluate the effectiveness of design in the innovation process (objective b) and one to evaluate how actors connect with design-enabled innovation to increase efficiency and competitiveness (objective c). Similarly, user benefit and business impact (objective a) – in the context of the Call 'theory of change' described above - could be seen as contributing factors to the over-arching expected goal of efficiency and competitiveness (objective c).

To address these aspects of ambiguity and overlap, the proposed methodology to support the impact evaluation of design-enabled innovation set out below takes a holistic approach, aiming to provide a common evaluation framework that can be applied to assess design-enabled innovation across any sector, at all levels – from the organisational to the European – and from the perspective of different actors, including SMEs and public administrations.

In order to support the specific requirements of the Co-Creation 02-2016 Call, this holistic framework is then applied to three 'application areas'. Each of these application areas is linked to a particular Call objective; each focuses on a specific evaluation question, and each is illustrated and exemplified by the evaluation of particular aspects of the Designscapes project, as shown in Table 1.

Application area	**Evaluation question**	**Call Objective addressed**	**Example from the Designscapes evaluation**
1. User benefits and business impacts	What user benefits and business impacts are associated with design-related policies and programmes?	a)	Calculating the economic and social benefits of the funding provided to projects by Designscapes using Cost Consequence Analysis (CCA)
2.Relationship between design and innovation	What contribution does design thinking make to innovation?	b)	Using Regression Analysis to predict the likely effects of the application of design thinking methods and tools to increase innovation
3.Value-creating networks and efficiency and competition	How does DEI support value-creating networks and in what ways do they increase efficiency and competitiveness?	c)	The contribution of the Designscapes financial instrument to creating value-driven networks and their impact on the efficiency and competitiveness of funded projects

Table 1: Common Impact Methodology application areas

For each application area the Common Evaluation Methodology provides

- methods and tools to apply the methodology in practice

- indicators to assess outcomes and impacts at different levels of analysis
- examples of evaluation results, as evidenced from the evaluation of the Designscapes project.

Our starting point is the over-arching Common Impact Evaluation Framework, which is described in Part A below.

PART A: COMMON IMPACT EVALUATION FRAMEWORK

Overall approach

The Designscapes Common Impact Evaluation Framework is based on a 'theory-driven' approach [2], an approach grounded in the 'realist' school of philosophy which looks at how something is supposed to work, with the goal of finding out what strategies work for which people, in what circumstances, and how [3]. Theory-driven evaluation emphasises the importance of taking context into consideration when assessing interventions, and allows for the perspectives and voices of different stakeholders to be considered – especially those whose voices are often not heard [4]. A theory-driven approach is essentially about testing a theory about what 'might cause change', even though that theory may not be explicit. One of the tasks of evaluation is therefore to make the theories within an intervention explicit, by developing clear hypotheses about how, and for whom, the intervention might 'work'. The implementation of the intervention, and the evaluation of it, then tests those hypotheses. This means collecting data, not just about intervention impacts, but also the processes of the intervention implementation, as well as data about the specific 'mechanisms' that might be creating change.

Mechanisms can be defined as: 'underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest' [5]. They are the combinations of 'resources' (the raw materials available to an intervention – for example its funding, the skills of the participants) and 'reasoning' (changes in the values, beliefs, attitudes, knowledge and behaviours of people and systems involved in the intervention) that, together, contribute to its outcomes and impacts. The way the mechanisms work depends on the 'context' in which they operate. An intervention will work – or not – in different ways for different people depending on 'contextual factors' – like the time and economic resources available to the people involved. There is always an interaction between context and mechanism, and that interaction is what creates the intervention's impacts or outcomes: Context + Mechanism = Outcome.

Why a theory-driven approach?

Theory-driven evaluation balances the need for an 'objective' and 'robust' assessment of an intervention's impacts with a recognition of the realities of working on the ground in complex and 'messy' situations . Ideally, project stakeholders – particularly those who fund it – look to the most robust evaluation approaches available in order to demonstrate results, impact and value. These approaches usually imply using 'experimental' methods to demonstrate results and impact – in

2 Chen, H-T (1990). *Theory*-driven *evaluations*. London: Sage Publications

3 Pawson, R. and N. Tilley (1997) Realistic Evaluation. London: SAGE

4 Guba E and Lincoln Y (1989). *Fourth Generation Evaluation.* London: Sage Publications

5 Astbury B and Leeuw F (2010) Unpacking black boxes: mechanisms and theory building in evaluation. American Journal of Evaluation 31(3): 363–81.

particular the use of 'Randomised controlled trials' (RCT's), which are seen as the 'gold standard' in evaluation and impacts assessment. [6] The attraction of experimental methods is that they are good at establishing the 'counterfactual'. Counterfactual evaluation involves comparing the outcomes of interest of those who have benefitted from an intervention (the 'treatment group') with those of a group similar in all respects to the treatment group (the 'comparison/control group'), but who have not been exposed to the intervention. The comparison group provides information on what would have happened to the participants in the intervention had they not been exposed to it.

However, a consistent problem with evaluation and impacts assessment in fields involving social interventions – as is the case with most design-led programmes and projects - is the difficulty in maintaining the 'temporal priority' required in RCTs - the assumption that a suspected cause precedes an event (for example, in clinical trials that the application of a particular drug will 'cause' the relief of particular symptoms). There are a number of factors that conspire to undermine temporal priority: history effects (the effects of 'external' variables unconnected with the intervention that may have an influence on its outcomes); selection effects (statistical bias in the treatment and control groups); instrumentation effects (using measurement tools in different settings); attrition (uneven loss of participating subjects). In short, the range and complexity of 'intervening variables' that may influence the effects of a social intervention are potentially unmanageable [7].

One way of handling these problems is to use quasi-experimental methods. These aim to reduce the negative influence of factors like history effects on the validity of the evaluation by compensating for the effects. One example is 'double difference' - comparing 'before and after' effects without randomisation. Another is using 'propensity scores' to statistically create comparable groups based on an analysis of the factors that influenced people's propensity to participate in the programme. 'Matched comparisons' involve matching individuals or groups who participate in a programme with similar individuals or groups who don't participate.

However, quasi-experimental methods themselves often prove inadequate for complex social interventions because they typically operate in turbulent, complex and unpredictable environments. In these situations, many evaluators argue that the only solution is to combine pragmatism – understanding and working with the characteristics and problematics of an intervention on the ground - with theory – building a model about the causal relationships that are assumed to make the

[6] Campbell, D.T. and J.C. Stanley, Experimental and Quasi-experimental designs for Research. Chicago, Rand-McNally, 1973.

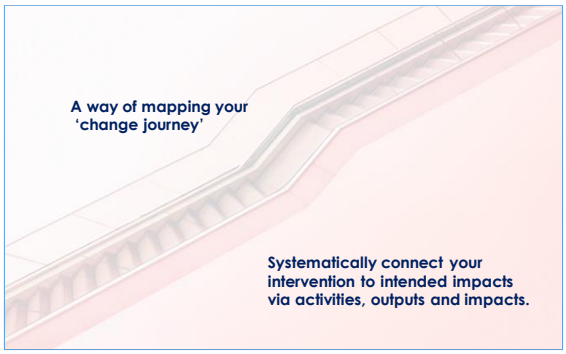
[7] Stromsdorfer, E.W. (1987).Evaluating CETA: Advances In Assessing Net Program Impact. Evaluation Review, 11:4.

intervention work (Befani, Barnett and Stern, 2014) [8]. This 'theory-driven' approach shapes the over-arching Common Impact Evaluation Framework described in the next section.

The Framework: Theory of Change

What is Theory of Change?

The framework of choice for putting theory-driven evaluation into practice in the field of design-enabled innovation is Theory of Change. As outlined above in the Introduction Theory of Change is a way of mapping the 'change journey' of an intervention so you can see the connections between the 'presenting problem' it wants to solve, the expected impact on that problem at the end of the intervention and everything that's supposed to happen in between. It doesn't matter much what the scale of the intervention is. It could be anything from a small design change an organisation wants to put into place to improve their efficiency to a major European-wide design-enabled innovation programme implemented through Structural Funds.



The essence of a theory of change: a way of mapping a change journey by breaking down a problem into components and linking them to create a logical pathway to a desired change.

The added value is that ToC incorporates an explicit 'theory' of what will cause that change. The job of the evaluation is to find out whether that theory works and how.

This change journey map starts with the ***presenting problem*** the intervention want to address. If we take the example of the Designscapes project, the presenting problem it wants to solve is:

> "many public sector organisations and businesses, especially SMEs, miss out on the potential to utilise design as a source for improving efficiency and stimulating growth".

It ends with the change the project wants to make to this problem after it has

[8] Befani, B, C Barnett and E Stern (2014). Rethinking Impact Evaluation for Development. IDS. Volume 45, Issue 6, pages 17–36, November 2014

completed its journey – in other words its ***expected impacts***. In the Designscapes case the longer term impacts the project hopes to achieve are increases in public and private sector efficiency and competitiveness.

To get from presenting problem to expected impacts, Designscapes carries out ***activities*** – for example developing technical and financial support instruments to stimulate design-led innovation projects.

These activities lead to the production of ***outputs*** such as a Designscapes' capacity-building programme, which helps funded projects to then produce their own design-led innovation outputs.

The utilization of these outputs lead to ***immediate outcomes*** (changes in awareness and increased knowledge), for example improved understandings of needs, demands, opportunities and challenges for design-enabled Innovations across Europe.

These immediate outcomes lead to ***intermediate outcomes*** (changes in behaviour and structures), for example SME's applying design-enabled innovations in their practice and scaling these innovations up and out.

Ultimately, these outcomes, combined together, will lead to the longer term ***impacts*** of public sector effectiveness and private sector competitiveness (Figure 2).

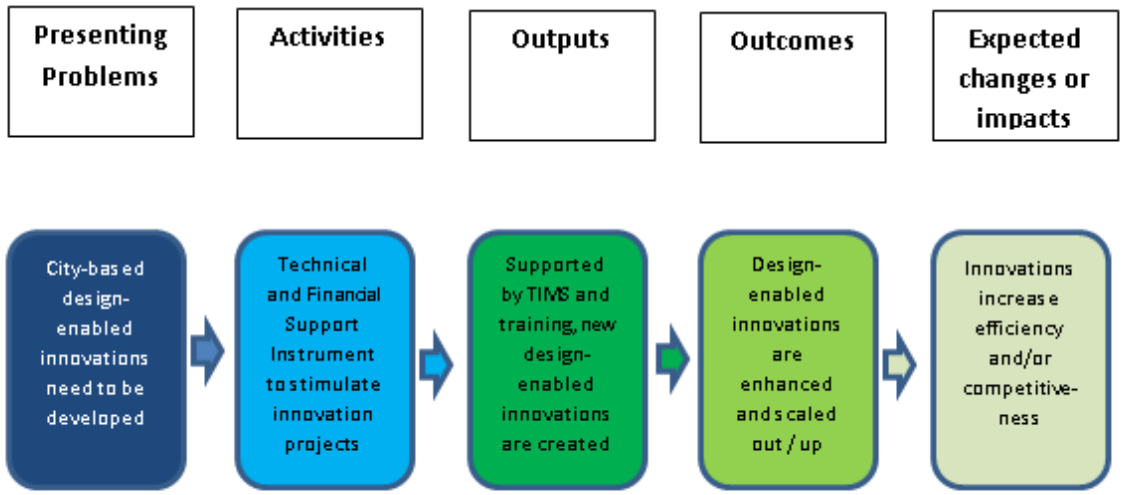


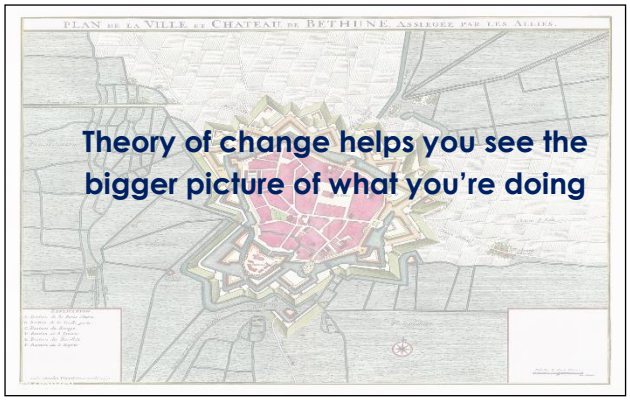
Figure 2: Simplified Theory of Change for the Designscapes project

But the value in Theory of Change is not just because it shows the logical pathway to a desired change. It goes beyond this 'intervention logic' to show the 'causal pathways' between a project's objectives, its activities, and its expected outcomes

and impacts. It says: "if we take action X, then this will cause effect Y and this will eventually lead to outcome Z". This makes it a powerful tool for impact evaluation.

Why use Theory of Change in DEI Evaluation?

Think of Theory of Change as the 'mothership' of evaluation. On one level it articulates the vision of an intervention as a whole – and the component parts that make up that vision. At a broader level it shows how the intervention and the problems it's trying to solve connects with the 'wicked problems' design-enabled innovation is grappling with in society as a whole.



Not only does Theory of Change provide a holistic view of the vision of the intervention and the interconnections that make up this vision, it situates that vision within the wider social context.

This helps to better articulate the contribution of the intervention to the broader societal impacts of DEI.

The capacity of Theory of Change to grasp the bigger picture means that it can provide a launch pad for a wide range of explorations in evaluation that can make a significant contribution to the success of an intervention. Theory of Change can be applied for different purposes (Figure 3). These include:

- Strategic / project planning
    - thinking through the link between what is being planned and the changes that are looking to achieved – this improves chances of success, by establishing the causal logic underpinning the intervention through developing 'if – then' statements
    - developing focused activities
    - producing a better allocation of resources
    - defining realistic outcomes that are clearly linked to activities
    - defining SMART indicators (Specific, Measurable, Achievable, Relevant and Time-bound) to understand impact at each stage of this journey forward
    - clarifying the conditions that need to be met to achieve success and the risks that need to be considered (assumptions)
    - producing a strategic plan and the actions needed to get there

- identifying and aligning key stakeholders that need to be involved

- Project implementation
  - Observing how the project is being delivered
  - Continuous monitoring of some key indicators provides evidence whether project is on track
  - Making adjustments in case of variation from plan
  - Assumptions can be tested, revised and activities corrected accordingly

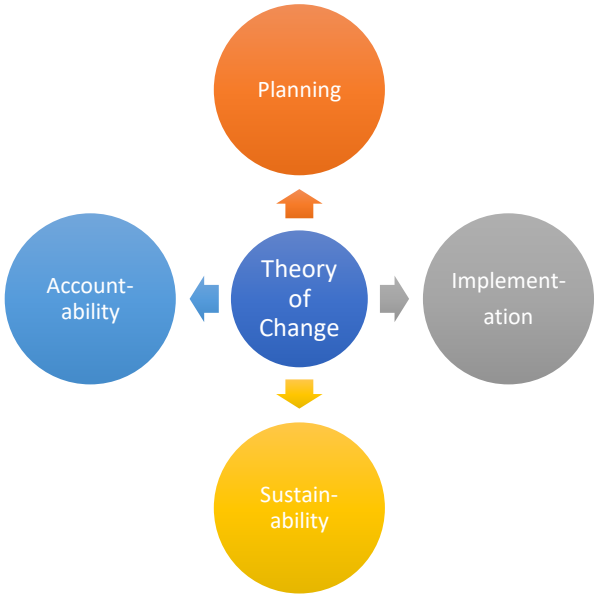


Figure 3: Purposes of Theory of Change

- Accountability
  - Using theory of change supports a more rigorous (scientific) approach to assessing the difference the intervention has made
  - It helps demonstrate the social impact because it allows evaluation to test the theory behind how the intervention was meant to work, whether and how this was realised in practice and if anything needs to change in the approach
  - It supports assessment of the social and economic value of the intervention, which can be used to demonstrate the value to funders of their investment
- Sustainability
  - Theory of Change generates evidence that the intervention works – this can be used to attract funders or sponsors
  - It supports calculating the replication potential of the intervention, and

the extent to which it can be scaled up and out

- It identifies those factors that are crucial for success.

And Theory of Change can be applied at different 'evaluation moments' through the intervention life cycle:

- At the design phase of the intervention ('ex-ante' evaluation), for example by organising co-creation workshops with stakeholders, using a Theory of Change, to help come to a shared understanding of key objectives and activities
- Over the implementation phase of the intervention ('process' evaluation), for example by helping to design a 'process dashboard' that shows the extent to which targets are being met
- At the end of the intervention ('summative' evaluation) – for example by comparing the expected outcomes in the Theory of Change with the actual outcomes identified through the evaluation evidence
- Post-intervention ('learning' evaluation) – for example by using the Theory of Change results to develop a sustainability plan for the intervention.

For the purposes of our Common Impact Evaluation Methodology, we are most interested in the 'accountability' purposes of Theory of Change and in its application in summative evaluation. It's important to bear in mind, however, that successful summative evaluation depends to a large extent on previously having used Theory of Change in the planning stage of an intervention , thereby establishing the 'causal logic' the summative evaluation will subsequently test. It's equally important to bear in mind that assessing the outcomes and impacts of an intervention at project end relies in part on data that have been previously collected through the project's 'process evaluation'.

Theory of Change specifies the underlying assumptions of a project and so incorporates a number of hypotheses about how the activities carried out as the project develops will *cause* changes at each stage of the project. The evaluation design and implementation approach follows this 'change journey'. The evaluation data collected along the way enables these embedded 'causal hypotheses' to be tested. If the evaluation data do not support a particular hypothesis, then this hypothesis needs to be discarded or modified. Theory of Change is therefore a powerful tool to assess ***'attribution'*** – whether the project outcomes and impacts that can be identified through evaluation can be attributed to the actions carried out by the project.

How can Theory of Change be used at different intervention scales and by different actors

In the context of impacts evaluation, another useful metaphor for understanding Theory of Change is to think of it as an architectural master-plan or blueprint. For

example transposing a trans-national policy on promoting design-enabled innovation across all EU member states to increase competitiveness into a Theory of Change Framework could provide an overview of all of the building blocks needed to achieve the desired impacts of that policy. These building blocks could then be narrowed down in scale to focus on a sub-area of the master-plan – for example a number of specific building types across different countries, or a particular building in a particular location – each of which would have its own Theory of Change.

In turn, the over-arching policy vision could be articulated through the lens of a particular perspective or point of view. So the owner of an SME in Poland might be interested in looking at the 'instrumental' benefits of design-enabled innovation – for example how DEI increases turnover - in the same way as a materials engineer might consider the architectural master-plan from the point of view of how DEI contributes to cost-savings as a result of increases in energy efficiency. These different perspectives imply different Theories of Change. In other words:

- Theory of change is a versatile and flexible tool that can be used at different levels of scale and in different contexts
- It can be adapted to reflect the perspectives and needs of different stakeholders
- Adaptation of the Theory of Change framework to different scales, contexts and stakeholder perspectives implies designing different kinds of assessment indicators and measurements that are appropriate for these different scales, contexts and stakeholder perspectives.

Table 2 overleaf provides illustrative examples of how Theory of Change could be used at different intervention scales in the DEI field.

Scale	Presenting Problem	Activities	Outputs	Outcomes	Impacts
Organisational (single actor)	Small atelier making customised designer-wear is experiencing mounting production costs and decreasing profit margins	Owners join EU-wide textile and clothing innovation network Set up exchange visits with network mentors Co-produce business innovation plan including use of laser cutting and 3D printing equipment, plus changes in shop floor production processes	Business innovation plan Installation of laser cutting and 3D printing equipment Re-organisation of production space	Production time per output unit decreased by 15% Production unit cost reduced by 23%	Turnover increased by 17% after 1 year Customer base increased 20% after 1 year Profit increased by 12% after 1 year
Project (partnership)	Across Europe, cities and towns are losing their cultural heritage artefacts. Not only is this a loss for the inhabitants	10 partners in 6 EU countries involved in using design thinking to explore factors affecting cultural heritage	Cultural Academy set up Cultural heritage preservation educational	Cultural Academy attracts over 300 members 1 year after launch Over 75% of	Cultural Academy model adopted by over 50 city administrations 2 years after launch

Scale	Presenting Problem	Activities	Outputs	Outcomes	Impacts
	but there is an increasing loss in cultural curation expertise	loss Co-creation workshops with stakeholders develop strategy for networking and educational events on cultural heritage preservation	programme delivers 10 international workshops and 10 online webinars, attended by 650 people in total Cultural Academy model disseminated to 4,000 stakeholders	educational programme participants report increased awareness of cultural heritage	Over 60% of city administrations report increased attendance at heritage events in their cities
Community/regional	Global food production is becoming more unsustainable and is a major contributor to climate change. Urban spaces in the region have untapped potential to develop innovative and	Research on feasibility of urban 'vertical farms', including use of AI and Industry 4.0 tools Co-creation workshops between stakeholder groups in cities across the	Vertical Farm Prototype 8 cities across the region take part in vertical farm piloting Pilot Evaluation Report Scalability Report	Vertical farm network increases awareness of urban sustainable farming across region 6 Public Administrations fund vertical farm experiments	5% of derelict urban properties rehabilitated to farming use in region Increase in regional food production of 10% 6% Reduction in

Scale	Presenting Problem	Activities	Outputs	Outcomes	Impacts
	sustainable farming methods	region Develop vertical farm prototype Pilot and evaluate prototype across region		5,000 customers purchase produce from vertical farms in 6 month trial period	food waste across region
National	Rates of unemployment of low skilled young people are increasing across the country in all cities. Job offers are unevenly distributed and there is a need to match jobs to young people	Develop a national jobs database co-ordinated nationwide Technical specification for a cloud-based technical platform and App to match young people with jobs Develop training programme for	National jobs database for young people Cloud-based technical platform and App to match young people with jobs 6,000 employment service staff participate in Training programme for employment	6,000 employment service staff improve digital competences and acquire job matching specialist skills Over 100,000 young people use the service 55% of users placed in jobs after 1 year of programme	Reduction in youth unemployed rate of 15% nationally Average annual saving of 13m euro on state welfare service costs

Scale	Presenting Problem	Activities	Outputs	Outcomes	Impacts
		employment service personnel to acquire digital skills to operate platform	service staff	operation	
Trans-national	The EU needs to become more competitive by capitalising untapped sources of growth through co-creation	EC launches 5 year programme to support design-enabled innovation through funding pilot projects in vulnerable economic sectors and for public administrations	Design-enabled Innovation funding instrument (DEIFI) launched through ESF Average 15 applications for funding each across all member states Evaluation framework and implementation plan set up	Participating entities become more aware of benefits of design thinking Participating organisations acquire design-led innovation skills through training and mentoring programmes Participating organisations introduce innovations into their practice	Aggregate increase in innovation across EU as measured by change in rate of patents filed EU-wide survey of companies and public administrations using behavioural additionality analysis shows net increase in efficiency of 12% Control-comparison

Scale	Presenting Problem	Activities	Outputs	Outcomes	Impacts
					analysis of outputs data in matched firms across EU shows 15 point increase in competitiveness

Table 2: Use of Theory of Change at different intervention scales

Table 3 shows how Theory of Change could be used by different actors in the DEI field.

Actor	Presenting Problem	Activities	Outputs	Outcomes	Impacts
SME	A small specialist furniture maker is experiencing falling demand due to high costs and changing consumer attitudes	Apply design thinking to 'outside the box' business re-structuring strategies Research Circular Economy business models Co-creation workshops with potential partners – property developers, landlords, tenants Develop and test new business model	Circular Economy furniture partnership Prototype Circular Economy furniture service for landlords and tenants	Increase in landlord and tenant awareness of benefits of Circular Economy Reduction in volume of furniture thrown away	Contribution to reduction in use of landfill sites Contribution to reducing poverty level of vulnerable people in rented property

Actor	Presenting Problem	Activities	Outputs	Outcomes	Impacts
Public Administration	Many people in many towns and cities experience loneliness and isolation. Public administrations have a duty to support them but they lack resources	Research on drivers of loneliness and isolation Co-creation workshops between public administrations and citizens Design platform and App to help neighbours connect with each other	Loneliness needs analysis report Anti-isolation network set up involving 80 public administrations Secure platform and App developed	70% of public administrations actively participate in the network Over 2,000 people use the platform and App in 6 month period 60% of users have connected with at least one neighbour	A survey of residents in the 80 participating public administrations shows self-reported loneliness and isolation levels have reduced by 23%

Table 3: Use of Theory of Change by different actors

How to develop a Theory of Change

There are seven steps to constructing a Theory of Change, as shown in Figure 4.

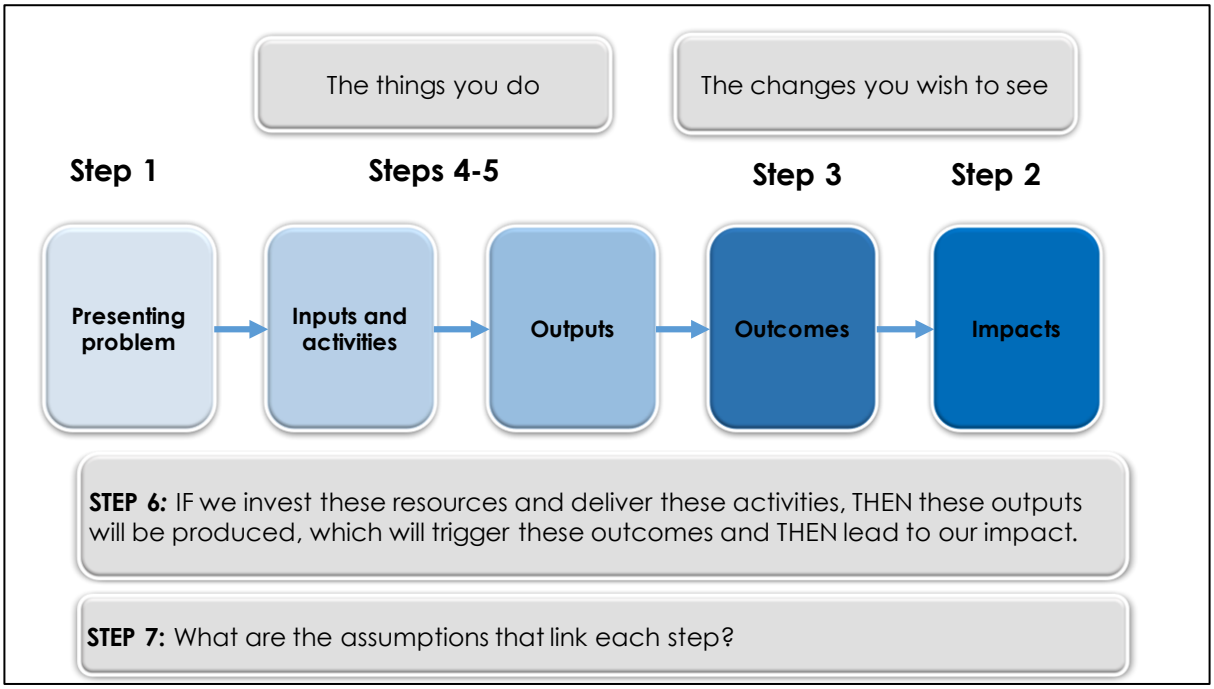


Figure 4: Steps in constructing a Theory of Change

- Step 1: Identify the problem the project is addressing and its underlying causes (the 'theory' of the problem)
- Step 2: Identify the long term project aim (expected impact) as well as more specific aims that link to this long term aim – i.e. the expected changes the project hopes to make to the presenting problem
- Step 3: Working backwards from the long-term aims, map the expected outcomes that lead to these long term aims.
- Step 4: Identify the inputs available (the resources - for example the funding, skills and so on) and the activities the project will carry out using these resources
- Step 5: identify the outputs these activities will deliver, that will in turn lead to the expected outcomes.
- Step 6: identify the causal connections that link the steps (the 'if-thens' that describe your theory of what will cause change)
- Step 7: identify the assumptions that need to be met if these causal connections are to lead to expected results.

A simple Theory of Change Mapping template is a good starting point to construct the Theory of Change. An example is provided below (Figure 5). This is taken from a project funded under the Designscapes project that aimed to use design thinking to

develop a new way of getting highly marginalised young people from an area in South London, UK to co-create solutions to fixing problems in their community through a 'Community Lab'.

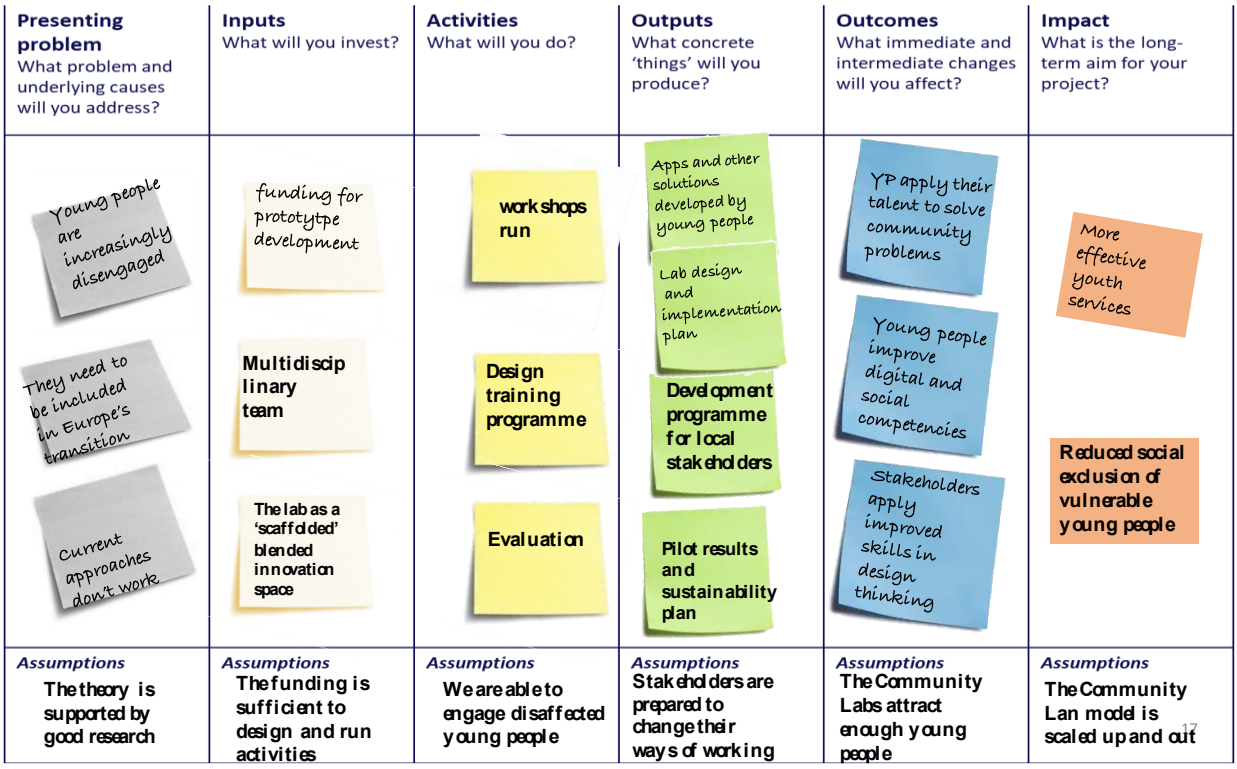


Figure 5: Theory of Change Mapping template for the Community Lab project

The Template maps out:

- The presenting problem – the issues the Community Lab wants to solve - and a theory of what causes problem - young people on the margins are being failed by the youth services system, which cannot effectively engage with them and broaden their life chances; they feel their horizons are limited and their prospects are poor; youth services have been severely cut in a sustained period of 'financial crisis' whilst at the same time are having to meet increasing demand by young people presenting with increasingly complex issues
- The inputs – the resources that are invested – an EC grant, co-funding; staff time, partner organisations, stakeholders; physical spaces to host the Community Labs
- The activities implemented – lifeworld analysis involving young people in exploring their needs; designing an interactive learning programme; action research projects
- The outputs produced by these activities – an interactive learning programme; co-designed action research experiments

- The immediate outcomes - changes in values, attitudes, knowledge, skills – that result from applying these outputs, for example Increased awareness by young people of digital tools
- The intermediate outcomes - changes in behaviours – that result from changes in knowledge and skills, for example young people more actively involved in community life
- Assumptions - beliefs about how a project will work, the people involved, and the external context, for example the environmental conditions are conducive to change; the resources invested are sufficient to carry out the planned activities.

The Mapping Template for the Community Lab goes a long way towards helping you collect and assemble the information needed to construct a Theory of Change. It takes you through most of the steps – from Step 1 to Step 5 – and explores the assumptions that underpin the project theory – Step 7. But the template doesn't capture the causal connections between the inputs, activities and so on. These are sometimes quite complicated because there is not always a direct linear relationship between an activity, an output and an outcome. One particular activity in a project may be linked to several outcomes. Some activities may be linked together. Some outcomes may be linked together.

An additional step in constructing a Theory of Change is therefore to produce a visual representation of the causal connections that together constitute the 'theory' of the project as a whole. An example for the Community Lab project is shown below (Figure 6).

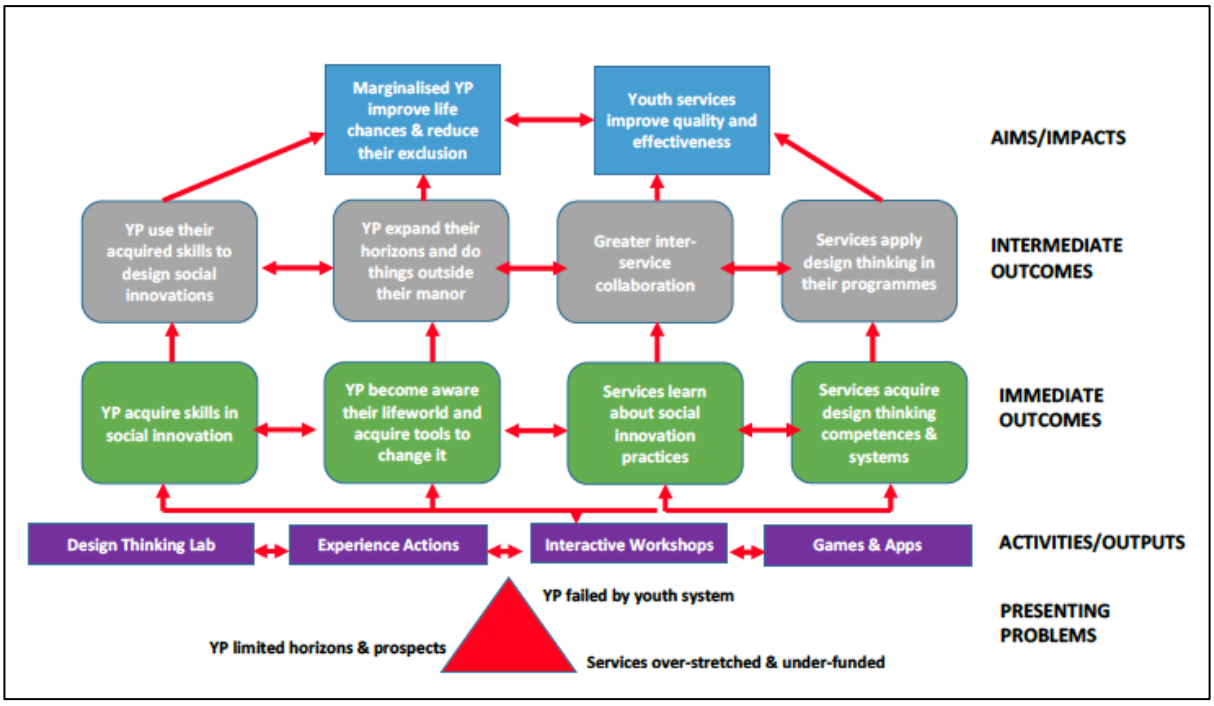


Figure 6: Creating a visual Theory of Change map

How to apply a Theory of Change in practice

Constructing a Theory of Change is only the starting point for DEI impact evaluation. The Theory of Change provides a framework and the parameters for the evaluation. The next step is to put that framework into practice by designing an implementation plan.

How to design an Impact Evaluation Implementation Plan

An impact evaluation implementation plan follows a clear progression – just as the intervention being evaluated – and its 'change journey' - does. The key stages in the life cycle of an impact evaluation are:

- Stage 1: Mapping and planning
- Stage 2: Implementation
- Stage 3: Analysis, Reporting and dissemination.

Each stage involves different activities that need to be considered, as shown in the checklist below.

Impact Evaluation Implementation Checklist: things to think about when developing an implementation plan

Stage 1: Mapping and Planning

This stage needs to identify:

- What are the purposes of the evaluation?
- Which main audiences or 'stakeholders will be interested in the results?
- How does the evaluation fit into the project 'life cycle'?
- What are the evaluation questions?
- What kinds of data collection and analysis methods and techniques will be suitable?
- How will the evaluation be integrated into the overall project plan?
- What role should the evaluator play (e.g. independent; participatory)?
- What involvement should users have in the evaluation process?

The main output of this stage will be an evaluation plan that:

- Defines the priority areas.
- Specifies the key evaluation questions that the evaluation will answer.
- Defines the methods and tools to be used to answer these questions.

- Shows how the timing of evaluation outputs will inform the key decisions of the project.
- Indicates the mechanisms and procedures to ensure feedback to all stakeholders
- Breaks down evaluation activities into its component activities, allocate responsibilities and make clear how the evaluation will be organised.

Stage 2: Implementation

Stage 2 involves putting the implementation plan into practice. The key steps are:

- Establishing the evaluation criteria that need to be assessed
- Deciding on what methods and techniques are to be used for data capture
- Managing and co-ordinating data collection, including analysing the results.

The initial exploratory and planning stage will have identified the sorts of questions different stakeholders want to ask. These questions need to be converted or translated into evaluation criteria to enable indicators to be defined and measurable data to be collected. There are no hard and fast rules on how to formulate evaluation criteria. The main influences on determining which criteria to specify are the 'object' and 'purposes' of the evaluation. Examples of evaluation criteria are things like 'usability', 'effectiveness', value for money'.

As with evaluation criteria, the selection of methods and techniques will be highly dependent on the ***object and purposes*** of evaluation. In this respect, some broad 'rules of thumb' to consider are:

- Impact evaluations, that consider changes retrospectively, for example, the ways in which target groups responded to an intervention, and in what ways their behaviours changed, typically utilise questionnaire surveys, interviews and focus groups
- In contrast, observation, diaries and logs are normally used in real-time as the intervention develops, and the cumulative evidence obtained then feeds into the summative evaluation at the end of the project
- The choice of particular data collection and analysis methods and tools depends on the complexity of the intervention. This in turn depends on the scale of the intervention and the perspective of the intervention the evaluation needs to capture. For example, a trans-European design-enabled innovation policy will require a broader set of evaluation methods and tools – to allow different perspectives to be compared against each other – than an evaluation of the effects of a design change implemented in an SME. This point is covered in more detail below.
- the selection of particular methods and techniques has skill and data resource implications. Some techniques (for example ethnographic methods) involve lengthy, in-depth field work producing copious amounts of data

- equally the selection of particular methods and techniques implies also using the appropriate type of data analysis (which has its own resource and skills implications).

The main output of this stage will be:

- An Evaluation toolkit, including overall methodology, indicators, list of methods and tools to be used in data collection and analysis, instruments for data collection
- Guidelines on how to use the instruments.

Stage 3: Analysis, Reporting and dissemination

This stage involves:

- Analysis of the data collected during the implementation phase
- Reviewing the results – for example comparing the analysis derived from different data collection methods to identify similarities and differences
- Integrating the results to produce evaluation conclusions
- Producing a report on the evaluation results, together with implications and recommendations as appropriate
- Ensuring the results are disseminated to maximise the learning derived from them.

Key issues to consider are:

- As noted above, the use of particular data collection methods and tools has implications for analysis. In general, large data sets (such as derived from surveys) normally need statistical software systems. Interpretative data (derived for example from observations) can be either analysed 'by hand' (for example using content analysis) or with proprietary qualitative software packages. Either way, the analysis is normally time and resource-intensive and requires good analytical and interpretation skills.
- The analysis needs to ensure that the evaluation is objective and unbiased (i.e. it reflects the voices of the stakeholders involved). In practice, evaluators try to collect and analyse a combination of different sets of data and compare these against each other. This is known as 'triangulation'. Triangulation makes it easier to identify 'causal pathways' between activities in an intervention and their effects. Triangulation entails the collection, analysis and synthesis of evidence of different types and from different sources, drawn from different kinds of evaluation activities, in order to arrive at conclusions in situations where attributing causality is difficult. In particular, a key aim of triangulation is to capture and reflect the 'voice' of different stakeholders in order to identify and understand their different positions and perspectives. Subsequent reporting of the results of the analysis needs to reflect these different voices.
- At the same time, reporting of evaluation results – and their dissemination - should be consistent with the 'purposes' of the evaluation. In other words, evaluations should be designed in terms of the decisions and actions they will inform. It is not

always easy to reflect this in recommendations, especially when the relevance of such recommendations may not be easily recognised by some stakeholders. The art of making useful recommendations lies in:

  - understanding the context in which the evaluation audience operates
  - addressing future realities rather than dwelling on the past
  - clarifying choices based on realistic options
  - showing how in practice recommendations can be implemented.

- Different stakeholders may require different communication and dissemination approaches. These might include:
  - Short summaries of the evaluation, tailored to different audiences
  - Journal articles for other researchers
  - Topical articles in the trade press (e.g. design magazines)
  - Workshops for specific audiences
  - Feedback co-creation workshops for key decision-makers.

Choosing the right methods, tools and measurements

Designing indicators

An impact evaluation is only as good as the data it uses. As noted in the preceding section, it's crucial that the evaluation is measuring what it's supposed to be measuring, and that the methods and tools it uses to collect those measurements are the right ones – i.e. they are appropriate for the purposes of the evaluation and its audiences.

Which comes first – the method or the measurement? In Theory of Change, the convention is to decide on the measurements – i.e. the indicators you need to use to assess whether the expected results developed in the Theory of Change Template and Map have been realised – first. Then you need to decide on the 'Means of Verification' of those indicators – the tools and processes needed to collect the data needed to measure these results.

Three things are typically measured in an impact assessment. These are:

- Outputs – the things produced by the activities of an intervention
- Outcomes – the changes associated with the use of these outputs
- Impacts – the longer term changes that result from the accumulation of outcomes over time.

These are described in the Box below.

Definitions of outputs, outcomes and impacts indicators

Output indicators relate to activity. They are typically measured in physical or monetary units (e.g., number of young people participating in a Community Lab.).

Outcome indicators relate firstly to the direct and ***immediate*** effect on beneficiaries brought about by an intervention (for example changes in the design skills of young people participating in a Community Lab). Secondly, they provide information on ***intermediate*** changes to the behaviour, capacity or performance of beneficiaries, organisations and systems involved in an intervention (for example an increase in the proportion of young people looking for employment after participating in the Community Lab).

Impact indicators refer to the consequences and broader and longer-term social and economic changes of the intervention beyond the immediate effects. Two concepts of impact can be defined: ***Specific*** impacts are those effects occurring after a certain lapse of time but which are, nonetheless, directly linked to the action taken and the direct beneficiaries (e.g. reduction in rates of anti-social behaviour in the area in which the Community Lab operates). ***Global*** impacts are longer-term effects affecting a wider population (e.g. increase in public sector efficiency rates in EU countries adopting design-enabled national policies)

Examples of these different categories of indicators in the DEI field are provided in Table 4, illustrated by the case of the Community Lab example described above.

Outputs	Immediate outcomes	Intermediate outcomes	Specific Impacts	Global Impacts
5 Community Labs in 5 EU cities 50 stakeholder organisations – including youth organisations join the project 100 marginalisd young people take part in the training programme 20 action research programmes co-produced in the Labs to solve	Increase in skills in using digital and media competence score (target: at least 50% of participating young people) following completion of training Increased knowledge of jobs in the creative and media sector (target: at least 50% of participating young people)	Young people have taken steps to get a job in the creative and media sector within 6 months of completing programme (target: at least 50% of participating young people) Participating stakeholder organisations increase service take up within 1 year of completing	Reduction in rates of anti-social behaviour in the areas participating in the project within 3 years after project end (target: 5%) Increase in youth employment in the areas participating in the project within 3 years after project	Savings in social care costs Increase in local fiscal revenue generated by increased youth employment

community problems	following completion of training increased understanding of using design thinking tools to work with young people (target: at least 50% of participating youth workers) following completion of training	programme (target: at least 50% of participating youth organisations) Problem-solving projects implemented in communities before project end(target: at least 10)	end (target: 5%)	

Table 4: Illustrative example of DEI indicator categories and types, Community Lab example

It's important to bear in mind that the indicators provided in the illustrative example above reflect a simple case of DEI. Note that the 'Global Impacts' indicators are unlikely to be realised in practice. This is because the scope, scale and duration of the Community Lab example - it's implemented in a relatively small geographical location, involving a small group of targeted beneficiaries over a 2-year period - mean that the intervention is unlikely to have an effect at the global level (although it's possible for the evaluation to say something about the potential contribution the intervention might make to global impacts). More complex cases – for example a major programme to stimulate competitiveness across EU countries – require a wider range of indicators measuring different aspects of the programme at different levels and in different time frames.

This highlights an important principle in indicators design. They need to fit the scope, scale, vision and objectives of the intervention. There are no hard and fast rules to help you to get it right every time. However, one heuristic applied in evaluation that is often used to help come up with a practical indicators solution is the 'S.M.A.R.T' guideline. This emphasises designing indicators that are:

- Specific (to the change being measured)
- Measurable (and unambiguous)
- Attainable (and sensitive)
- Relevant (and easy to collect)
- Time bound (with term dates for measurement).

Table 5 shows illustrative examples of outcomes measures for the Community Lab

project that fit the SMART criteria compared with those that don't fit the criteria.

Criteria	Fit the criteria	Don't fit the criteria
Specific	Increase in digital and media competence score on the DigComp Index	Increase in digital skills
Measurable	Change in participant self-reported score on Bristol social inclusion matrix	Increased social inclusion
Attainable	Completion of pass mark in digital skills training module	Passing Oxford University Entrance examination
Relevant	No. of problem-solving projects implemented in communities	No. of communities signing up to EU town twinning programme
Time bound	% participating young people who have taken steps to get a job in the creative and media sector within 6 months of completing programme	% young people getting job in creative and media sector

Table 5: Examples of use of SMART criteria for outcomes indicators

However, there are some issues that need to be borne in mind when using SMART criteria to develop indicators in the DEI field. These are:

- SMART indicators prioritise the collection of quantitative data. Numbers are good at providing a picture of the magnitude of an effect on the surface – but they are not so good a revealing what's happening in depth
- SMART indicators tend to be imposed from above – they reflect the 'expert' perspective rather than the lived experience of people on the ground
- Because of this, SMART indicators are often not 'co-created' – they don't involve evaluators and 'experts' working together with the 'subjects' of the intervention.

Although, as noted above, it's a good idea to take into account who the audience is for the evaluation when designing it, the preferences of stakeholders – particularly funders - tends to be for quantitative data (because 'numbers' and 'statistics' are generally seen as more robust and 'objective'). Whilst quantitative indicators, as the name suggests, are useful for measuring quantity – how much change has occurred - they're less useful at measuring the quality of change. For example, the number of

stakeholders who get involved in a Community Lab is less important than the quality of the relationships that they form when collaborating together. Qualitative indicators that capture the stories and narratives of the beneficiaries of an intervention are critical to understanding how it works.

This links to the two other issues cited above. Too much emphasis on quantifying the impact of an intervention, in line with the expectations of more powerful stakeholders, risks imposing from above a vision of what change means onto people who's lives are affected by the intervention, but who don't have a voice in its evaluation. There is therefore a strong case to be made for working with stakeholders – in particular the target groups and communities in which the intervention is implemented – to co-produce assessment indicators. Indeed, co-creation in evaluation is consistent with the key principles of design thinking, as set out in the Box below.

Principles of Design Thinking (Gobble, 2014; IDEO, 2014) [9] [10]

- human-centred and participatory – the needs and wants of the users and all the people who will be affected by the design project should be at the core of the design process
- multiplicity – design thinking embraces 'multiple realities'. It encourages 'divergent thinking' and 'thinking outside the box', whilst at the same time supporting a collective and shared vision through convergent thinking
- societal focus – design thinking aims to explore new ways of addressing the persistent and intractable 'wicked problems' like climate change, inequality and economic unsustainability that threaten our survival (Rittell and Webber, 1973) [11]
- co-creation – design thinking aims to produce 'co-created' solutions to wicked problems by actively engaging the disengaged in policy and practice by surfacing 'out of the box' and 'disruptive' ideas to address entrenched problems
- empathising – the starting point for co-creation is empathising - gaining an empathetic understanding of the problem by getting immersed in the 'lived experience' of those most affected by the problem. Empathising provides the foundation for the subsequent stages of design thinking: defining – stating the problem from a human perspective; ideating – identifying new solutions by thinking outside the box; prototyping – developing new solutions to the problem; testing – evaluating the solutions.

[9] Gobble, M (2014) Design Thinking, Research-Technology Management, 57:3, 59-62

[10] https://www.ideo.com/post/design-kit

[11] Rittel, H W and Webber, M M (1973). Dilemmas in a General Theory of Planning. Policy Sciences, Vol. 4, No. 2

Bringing co-creation into indicators design – for example through running 'structured co-creation workshops' in the evaluation - has a number of benefits for the evaluation, including:

- it results in more realistic, meaningful and achievable indicators than those set by top-down methods
- it highlights the different information needs and ideas of change of different stakeholders and community groups – a good example being differences associated with gender, which is almost never referred to in evaluation
- it enables impact assessment to focus not just on what is measured but on how it is measured and who has decided on what to measure
- it contributes to providing information on the why and how of change
- it helps create ownership and buy-in to an intervention and its evaluation – particularly from people who are often categorised as 'hard to reach'.

Adopting a co-creation approach to designing indicators doesn't mean exchanging the 'SMART' acronym for some inferior brand - note that the properties of co-created indicators cited in the first bullet point above refer to 'realistic, meaningful and achievable' indicators – attributes that are very close to the SMART approach. What's proposed is to take the best of SMART and add to it supplementary indicator design principles and techniques that focus on assessing quality of change, whilst embedding the process in a participatory design ethos. One way of doing this is the 'SPICED' guideline: [12]

- Subjective – emphasises the insights and 'lived experience' that people on the ground can bring to the evaluation
- Participatory – involving project beneficiaries as well as other stakeholders in a co-creation process
- Interpreted (and communicable) – indicators derived from on the ground lived experience may often need to be interpreted so as to make sense to evaluators and experts
- Cross-checked – speaks to the need for 'triangulation' of data and indicators from different sources and representing different perspectives that can then be compared against each other to produce a balanced result
- Empowering – the process of co-creation of indicators is empowering in itself and allows people with a stake in the intervention to critically reflect on their situation and how it needs to change

12 https://www.betterevaluation.org/sites/default/files/EA_PM%26E_toolkit_module_2_objectives%26indicators_for_publication.pdf

- Diverse and disaggregated – indicators should reflect the diversity of the lived experience of different groups.

Choosing the right methods and tools

Having made a decision on what is going to measured, the next step in impact evaluation is to decide on the best way of collecting the information that enables the measurement to take place. It almost goes without saying that the first principle of data collection methods and tools is: make sure they are aligned with the indicators being used to measure impact. For example, there's little point designing and delivering a beneficiary survey if you want to measure the cognitive changes attributable to use of a piece of learning technology equipment in real time.

The principle of 'indicator-method' alignment is not just a simple operational issue – choosing the methods and tools that will do the job on the ground. Selecting available methods and tools – or designing new ones – to collect the information that needs to be collected also involves 'normative', 'values-driven' and 'ideological' choices.

This goes back to our 'Theory of Change' – and its central premise that you need to specify a theory of the causes of the presenting problem you want to solve, as well as a theory of what will cause a change to that presenting problem. These theories will inevitably shape the indicators selected to assess whether changes can be identified and to which factors the changes can be attributed, and in turn will shape the choice of methods and tool to collect the evaluation data. In some cases these theories are implicit. They don't have a name. In other cases the theories are grounded in an established – and named - school of thought that incorporates theoretical positions and assumptions that will need to be tested through the selection of appropriate measurements and collection of relevant data. An example from the DEI world is shown in the Box below.

Urban Import Replacement

One of the conceptual frameworks that has influenced understandings of the impact of design-enabled innovation in an urban context has been the idea of 'urban import replacement'. First developed by Jane Jacobs in the late sixties 13, it argues that economic growth is driven by import replacement, which in turn happens when a city or urban region begins to locally produce goods it previously imported. A more recent application of the theory suggests urban import replacement works as a driver of growth through the emergence of ecosystems of venture capitalists, entrepreneurs, social innovators and public administrations working together not only within a discrete urban territory but as interconnected

13 Jacobs J (1969). *The Economy of Cities*. New York: Random House.

innovation engines across territories. In this way a diversified group of innovative and versatile cities can through trade learn from each other and not only grow but develop "on one another's shoulders" as Jacobs put it 14 (for a critique of the theory see David Harvey 15). Evaluations taking an urban import replacement theory position on the impact of a policy on design-enabled innovation would likely choose some measure of design-led innovation – for example incidence of multi-stakeholder innovation networks over time – and set this against a measure of import replacement – for example changes in the number of electronic bicycles produced over a period in the territory under investigation. The methods used to make this assessment would likely involve statistical analysis of economic datasets.

The choice of measurements, methods and tools can also be shaped by a theoretical perspective that is not directly 'domain-driven' – as in the case of urban import replacement theory – but by a particular methodological 'stance' the evaluation is taking. We began this Guideline on how to do impact evaluation in DEI with an explicit proposal that the best methodological approach to use is 'theory-driven' evaluation. 'Theory-driven' evaluation takes a particular ontological position – assumptions about what things exist in the social world and the nature of social reality - and a particular epistemological position – assumptions about the nature of knowledge and how we come to know social reality. These assumptions can be summarised as follows:

- Social programmes and interventions are viewed as an attempt to address an existing social problem – that is, to create some level of social change. The focus of evaluation should therefore be on assessing whether and how this change has occurred.
- Programmes and interventions work by enabling participants to make different choices, so a key objective of evaluation is to capture how and why these choices are made.
- Making and sustaining different choices requires a change in participant's 'reasoning' (for example, values, beliefs, attitudes, or the logic they apply to a particular situation) and the resources (e.g. information, skills, material resources, support) they have available to them. This combination of 'reasoning and resources' is what enables the programme to 'work' and is defined as a program 'mechanism'.
- Programmes and interventions work in different ways for different people - a key task of evaluation is therefore to capture 'what works, for whom under what conditions'

14 Flaccavento, A (2016). Building a Healthy Economy from the Bottom Up. Lexington: University Press of Kentucky

15 Harvey, D (2013). Rebel Cities: From the Right to the City to the Urban Revolution. Verso: London

- A key task of evaluation is to make sure that the voice of all stakeholders is reflected in the evaluation – particularly those who have less power, for example those who tend to be marginalised by 'the system'. This means the use of methods and tools that can capture the experience of those on the margins.

As noted above, the methodological stance evaluation takes should speak to the purposes of the evaluation and the audiences it is aimed at. Some would argue that the main purposes of evaluation are objectivity and accountability – and the key audiences are whoever has the biggest 'stake' in the evaluation – which is often taken to be policy-makers and funders. From this perspective, impact evaluation should always strive for the most objective and robust 'gold standard' methodological stance, i.e. experimental methods using RCTs.

However, there is a counter-argument that a methodological stance that is grounded in 'theory-driven' and 'realist' evaluation is more in line and consistent with the ontological and epistemological principles of design thinking, as set out in the preceding section i.e.: human-centred and participatory; supporting a collective and shared vision through convergent thinking; focusing on societal issues and 'wicked problems' like climate change, inequality and economic unsustainability; focusing on co-creation by actively engaging the disengaged in policy and practice.

One way of making sense of these different 'methodological stances' is to think about the 'research paradigms' that underpin them. As defined by Kuhn (1962) a paradigm is

> *"a world view, a general perspective, a way of breaking down the complexity of the real world. As such paradigms are deeply embedded in the socialisation of adherents and practitioners: paradigms tell them what is important, legitimate, and reasonable".* [16]

As this definition indicates, there are two elements to a paradigm: its content (in terms of theories, research methods and assumptions about the nature of the phenomena being studied) and the processes by which its adherents come to acquire these. In the evaluation field, the paradigms being employed can be hard to untangle. The research methods employed by practitioners in evaluation have been developed in a multitude of different 'schools' and traditions, which operate within sometimes widely varying research paradigms. These different research approaches can be grouped together and labelled in many different ways. A simple – if crude – categorisation that has long been adopted is between 'positivist' (or experimental) and 'non-positivist' (or non-experimental) paradigms (Halfpenny, 1982) [17]. Although to some extent there is a degree of clarity about what constitutes

16 Kuhn, T S (1962); *the structure of scientific revolutions.* Chicago: University of Chicago Press

17 Halfpenny, P (1982) *Positivism and sociology, explaining social life*, London: George Allen and Unwin

'experimental' research there is a wide variety of thinking about what constitutes 'non-experimental' research, ranging from 'constructivist'(e.g. Guba and Lincoln, 1989) through phenomenological, holistic, ethnographic and biographic (e.g. Denzin,1994), to post-modern research (e.g. Roseneau 1992) and so on.

Whatever label you want to put on them, research paradigms will affect the methodological stance of an evaluation, which will in turn affect the methods and tools used to collect data. It's useful in this context to think about impact evaluation as a process that works towards increasing specificity, as illustrated in Figure 7.

As Figure 7 shows, the methodological stance (or methodology for short) defines the over-arching direction the evaluation takes, and incorporates a particular research paradigm – the 'world view' or general perspective of the evaluation. This then influences the 'methods' used in the evaluation – the broad processes through which data are collected. Methods in turn shape the evaluation 'tools' used to collect the data. Tools are the specific instruments that collect specific types of data on the ground.

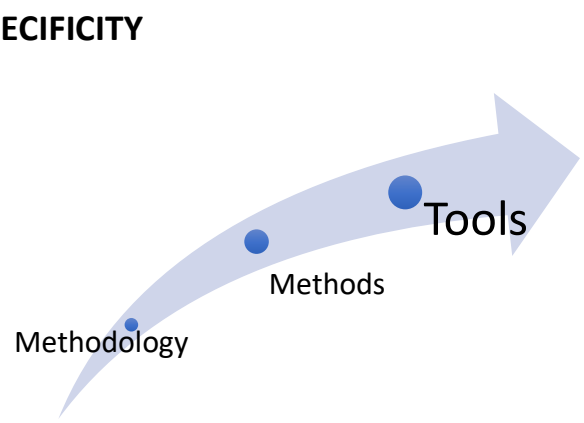


Figure 7:Relationship between evaluation Methodology, Methods and Tools

To avoid getting lost in the sometimes disorientating forest of different paradigms, let's illustrate the relationship between methodology, methods and tools by comparing an experimental methodology with an example of a non-experimental methodology – theory-driven evaluation (Table 6).

Methodology	Research paradigm	Method Examples	Tools
Experimental	Nomologist – aims to establish general regularities or systems to explain	Randomised Controlled Trials Pre-test/post-test	Wide range of instruments used to collect data to compare

Methodology	Research paradigm	Method Examples	Tools
	phenomena. What exists can only be established through empirical verification.	Instrumental variables Multiple Regression Difference-in-Differences Statistical Matching	treatment and comparison groups, depending on the assessment indicators of interest. Can use either 'secondary data' – like statistical datasets (e.g. GDP) – or primary data collected from the field, e.g. programme participation rates; surveys of programme participants
Theory-driven	Rationalist – aims to establish regularities and common structures but accepts the notion of 'multiple, socially constructed realities'. What exists is mediated through context	Surveys Field studies Interpretative Participatory Econometric	Questionnaires; interviews Observation; case studies; ethnography Content analysis; Discourse analysis Action research SROI; Cost Consequence Analysis

Table 6: Comparison between experimental and theory-driven impact evaluation methodologies

]It should be noted that Table 6 is intended to just give a flavour of the two methodological stances and to highlight their contrasting features. There is neither space nor scope in this brief Guideline to do justice to the many variations of methods and tools that are used in impact evaluation. For a comprehensive guide on using experimental methods in impact evaluation see the World Bank 's 'Impact

Evaluation in Practice, 2nd Edition' [18]. With regard to methods and tools used in theory-driven evaluation we present in Part 2 of this Guideline some examples of methods and tools used in the DEI domain.

One obvious difference between the two methodological stances highlighted in the Table is that the experimental approach essentially sees the core objective of impact evaluation as establishing the 'truth' about an intervention – whether it has made a difference and what causes that difference. It does this by providing 'counterfactual' evidence through comparing the effects of an intervention on those exposed to it – the 'treatment' group – with the situation of a comparable group who have not been exposed to it. In contrast, theory-driven evaluation sees the core objective of impact evaluation as understanding 'what works in the intervention, for whom and under what circumstances'. In other words, there is an acceptance that the reality of change may be different as a result of the intervention for different people, in different ways and because of different factors.

Can you mix and match methods and tools from different methodologies? The short answer is – it depends on your paradigm. A 'purist' perspective sees experimental and non-experimental methodologies as completely incompatible (Guba and Lincoln, 1989) 19. Yet, in practice, evaluations do mix and match different paradigms. Experimental methodologies will favour methods with a strong quantitative data analysis element, since these transform the information collected into data which can be relatively easily manipulated by statistical methods. But quantitative methods will often be supplemented by qualitative methods when these are appropriate, although the way in which these are used, and interpreted, tends to be based on the quantification of qualitative data through, for example, frequency counts of 'events'.

In the non experimental methodologies, both qualitative and quantitative research methods are also used, although these tend to favour qualitative methods which are seen as particularly appropriate for capturing the complexities of the individual situation and subjective experience.

Even Guba and Lincoln concede that some situations call for supplementing 'constructivist' methods with quantitative methods – although they dispute whether statistics can be used to infer causality:

> *'We have argued often that evaluators who operate in a constructivist, responsive and now fourth generation mode will use primarily although not exclusively qualitative methods. But there will be times when quantitative methods - tests or other measurement instruments or numeric displays - will be, and should be, used. The single limitation that a constructivist, responsive,*

[18] https://openknowledge.worldbank.org/handle/10986/25030

19 Guba E and Lincoln Y (1989.*) Fourth Generation Evaluation*. London: Sage Publications

fourth generation evaluator would put on the use of quantitative methods is that no causally inferential statistics would be employed since the causal linkage implied by such statistics are contrary to the position on causality that phenomenologically oriented and constructivist inquiry takes"(Guba and Lincoln, 1989).

If you don't take the line that experimental and theory-driven methodologies are incompatible, there are a number of ways in which they can co-exist within an impact evaluation strategy:

- Complementary model **-** different research approaches can be seen as different models for generating research data on a particular programme. Different models might be used, either at different stages in the development of the programme, or simultaneously; either within the boundary of one evaluation or within a parallel study. These strategies provide the opportunity to address different aspects of the programme through appropriate methods, and to meet the needs of different stakeholders, even if their expectations from the research are at variance from one another. For example, the requirements of those seeking 'scientific proof' can be met, at the same time as other data is collected that will meet the needs of those requiring a more in depth understanding of the intervention and its impact
- Simultaneous model - other research approaches would be used alongside the RCT in order to gather information which can be used to triangulate the conclusions derived from the RCT. For example an 'interpretative' approach might be used to gather an alternative set of information about the process of the intervention, and its effectiveness from the point of view of different participants. The experimental component of the evaluation – the RCT - would then focus on the 'simple' and 'consensus' elements of the intervention.
- Sequential model **- e**xploratory work – for example using ethnographic methods, participatory co-design workshops – is carried out to review key questions like: how far a consensus exists about the underlying cause of the problem to which the intervention is being targeted; are the conditions necessary for applying an experimental method likely to be met. An experimental methodology may or may not be subsequently applied depending on the results of this ground work.

Ultimately, choosing the right methods and tools for impact evaluation is a complicated business which involves a balancing act that trades off purposes, audiences, research paradigms and resources against each other in order to come to an optimal solution that delivers the best results. There are no hard and fast rules for doing this but the Box below provides some guidelines.

Summary: choosing the right methods and tools

- Make sure the methods and tools you choose are aligned and are compatible with the indicators selected for the evaluation and that these

methods and tools will be capable of collecting the data required to measure what needs too be measured

- Review your Theory of Change to highlight the underlying theories and assumptions that describe the presenting problem and the change you hope to make to that problem. These theories and assumptions should guide the methods and tools needed to carry out the evaluation
- Review the purposes of the evaluation and the audience the evaluation is aimed at. If the purposes of the evaluation are primarily to quantify change, and the audience is likely to want to see results that are backed up by statistics that infer the causes of that change, then consider using an experimental methodology
- Bear in mind however that 'classical' experimental methods – like randomised controlled trials - are unlikely to work in the kinds of complex social situations in which DEI operates. If the rigorous conditions needed to apply RCTs cannot be met, then use an alternative method – the most likely alternative being a 'theory-driven' methodology
- In any case, although it's important to consider the evaluation audience, you also need to think about choosing methods and tools that give a voice to other stakeholders – particularly those who often don't have power in an intervention – typically the target group and their communities. Giving a voice to the less powerful often means using qualitative methods and tools – for example ethnographic methods that capture the 'lived experience' of people through stories and anecdotes
- Choosing a methodology like theory-driven evaluation doesn't mean the methods and tools used have to be less 'rigorous'. It's perfectly acceptable to include quasi-experimental methods in a theory-driven evaluation – for example a longitudinal survey comparing a group of participants involved in the intervention with a similar group not involved
- More broadly, experimental methodologies can be combined with non-experimental methodologies in a 'mix and match' approach in which experimental and non-experimental methods are used to complement each other, in sequence or in parallel, depending on the characteristics of the intervention.

Putting it all together: integrating the results

The final stage in impact evaluation – prior to reporting on and disseminating its results (which is covered above in Section ?) – is to gather all of the information that has been collected and analysed using the methods outlined in the previous section in order to come to conclusions that are supported by the evidence.

This isn't as simple a job as it sounds – especially for complex interventions that have collected and analysed data using a range of methods and tools and which reflect the sometimes conflicting positions of different stakeholders.

Presented below are three methods to integrate evaluation results, starting from the simplest to the most complex method:

- Basic triangulation
- Theory of Change Analysis
- Contribution analysis.

Basic Triangulation

What is it

Triangulation allows for the synthesis of evidence of different types and from different sources, drawn from different kinds of evaluation activities, in order to arrive at conclusions. A key aim of triangulation is to capture and reflect the 'voice' of different stakeholders in order to identify and understand their different positions and perspectives. Triangulation is essential in a realist evaluation approach for the following reasons:

- First, it allows for the capture of complex contextual dat.
- Second, it avoids relying on 'expert' knowledge and evidence (for example that derived solely from peer-reviewed journals) and
- third, it provides a means to consider ideologies, values and power relations between different actors.

Triangulation supports generalisability and transferability of findings especially in a situation where the intervention is innovative and evolving, and the evidence base is limited. This is because it increases the 'robustness' and transferability of findings through cross-checking of data derived from different sources and from different actors thus helping to boost the internal validity of the research.[14] Triangulation entails integration and synthesis of the evaluation evidence from the different evaluation activities carried out in the intervention, combining secondary data (e.g. content analysis of reports produced by the intervention) and primary data (e.g. from beneficiary surveys); quantitative analysis with qualitative analysis; information reflecting the 'official' point of view – for example project managers - with information 'on the ground' – for example from community representatives. [20] [21]

[20] O' Donoghue, T. And Punch K. (2003). Qualitative Educational Research In Action: Doing And Reflecting. London: Routledge

[21] Patton M Q (1999). Enhancing the quality and credibility of qualitative analysis. Health Serv Res., 34(5 Pt 2):1189-208.

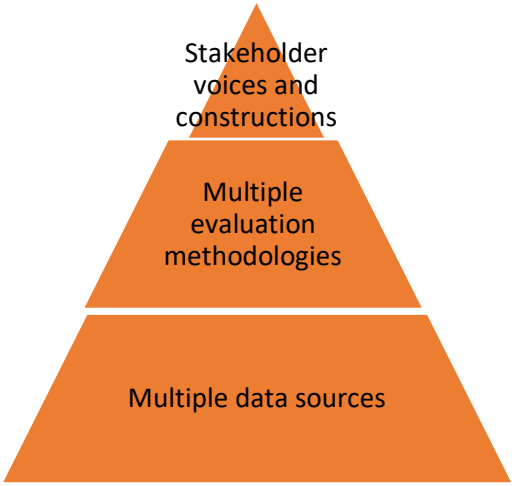


Figure 8:: Triangulation

Triangulation compares different sides of the same evaluation 'object' so as to come to a conclusion about the nature of that object.

This comparison can be done in several ways:

- by comparing the results of two or more evaluation methodologies (e.g. experimental and non-experimental)
- by comparing the results of different methods used in the evaluation (e.g. data from a beneficiary survey compared with data from observation)
- by comparing what different actors in the

When should it be used

Triangulation is a standard technique in impact evaluation. It should be used routinely when the impact evaluation involves the use of more than one data collection method and tool. Triangulation provides clarity to evaluation but has limitations. It's sometimes difficult to apply to data in a uniform and consistent way. In addition, it's possible that the findings of two or more data sources may be inconsistent or conflicting. For these reasons, the main value of triangulation is to add credibility to the evaluation conclusions and to provide reassurance that data analysis has been carried out systematically. Triangulation will provide a better grasp of the 'bigger picture' of the evaluation but it won't necessarily reveal the 'causal pathways' that lead to identified outcomes.

How to do it

A simple but effective way to triangulate evaluation data is to use an evaluation matrix that compares the results of the different methods and tools used in the evaluation set against the key evaluation questions the evaluation is intended to answer. An example is shown below for an intervention that uses design thinking to improve the efficiency and effectiveness of an organisation providing services for people dealing with substance misuse and their families (Figure 8). The Theory of Change developed for the intervention proposes a new service delivery model that has been developed through a series of structured co-creation workshops with clients, their families and service staff. The new service model focuses on a post-treatment rehabilitation strategy that is tailored to each individual client and is delivered through 'joined up' working between various services, including mental health, social services and welfare services. The hypothesis behind the intervention is that repeated drug use post-treatment is more effectively prevented through multi-service after care.

The impact evaluation design combines the following methods and tools:

- A longitudinal cohort study involving a treatment group – clients who have volunteered for the pilot multi-service after-care programme – and a control group – clients with a similar profile to the treatment group who are involved in a traditional after care programme. The use behaviours of both groups are monitored before the launch of the pilot - pre-test - and after the pilot has finished - post-test (quasi-experimental method).
- A user survey delivered to the treatment and control groups before the launch of the pilot and after the pilot has finished that collects self-reported information on client well-being (including the Warwick-Edinburgh Mental Well-Being Scale – WEMWB)
- User diaries completed by a representative sample of the treatment and control groups over the duration of the pilot that collect detailed information on the lived experience of clients over the pilot.

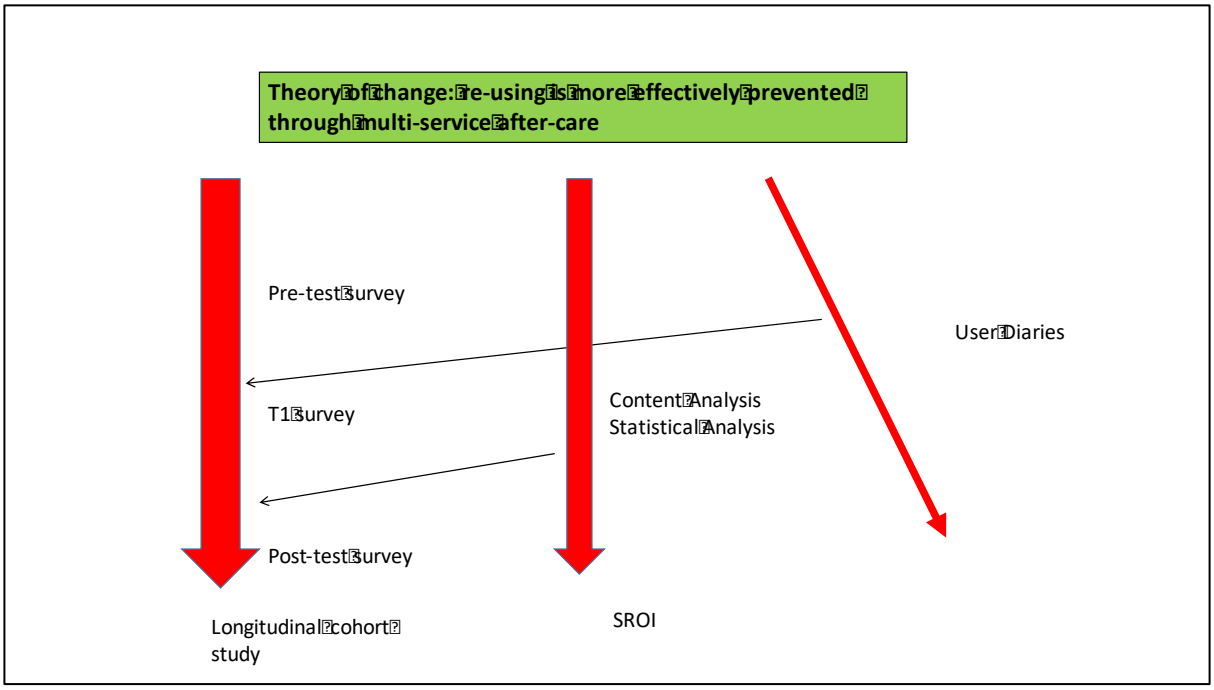


Figure 9: Triangulating data sources

The data collected through these methods and tools is analysed using appropriate analysis tools – including content analysis of the user diaries and statistical analysis of data from the cohort study and user surveys. Data from the cohort study and user survey is combined with economic data that models treatment costs together with other cost data (including proxy measures on estimated costs of continuing treatment and associated welfare costs) to produce an analysis of the Social Return on Investment (SROI) of the programme.

The results feed into an Evaluation Matrix (Table 7) which allows triangulation of the results set against the key evaluation questions.

Key Evaluation Questions/Method	Cohort Study	User Survey	User diaries	SROI
KEQ1: What has been the impact of the pilot programme on drug use behaviour?				
KEQ2: What additional personal benefits for clients and their families can be attributed to the pilot programme?				
KEQ3: Did the pilot programme create unexpected challenges for participants or generate unforeseen impacts?				
KEQ4: What is the economic return of the pilot programme compared with conventional programmes?				

Table 7: Evaluation Matrix

Using the Evaluation Matrix, the key results obtained for each of the evaluation methods and tools used in the impact evaluation are set out for each of the key evaluation questions. Each result specified is backed up by supporting evidence. The completed Matrix can then be used to provide an impact evaluation report together with appropriate conclusions and recommendations.

Theory of Change Analysis

What is it

As outlined above, an intervention's Theory of Change specifies the underlying assumptions of the intervention and so incorporates a number of hypotheses about how the activities carried out as the intervention develops will promote changes at each stage of its life cycle. In 'summative' evaluation mode, ToC analysis essentially compares the 'baseline' ToC established at the beginning of the intervention (focusing on the expected outputs and results identified in this baseline) with actual outputs and results to assess how far the intervention has travelled on its 'change journey'. It therefore assesses the 'distance travelled' - towards expected outcomes and impacts. This can also contribute to establishing a 'counterfactual' for the intervention – what would likely have happened, and what are the likely implications for project beneficiaries, if it had not been implemented

When should it be used

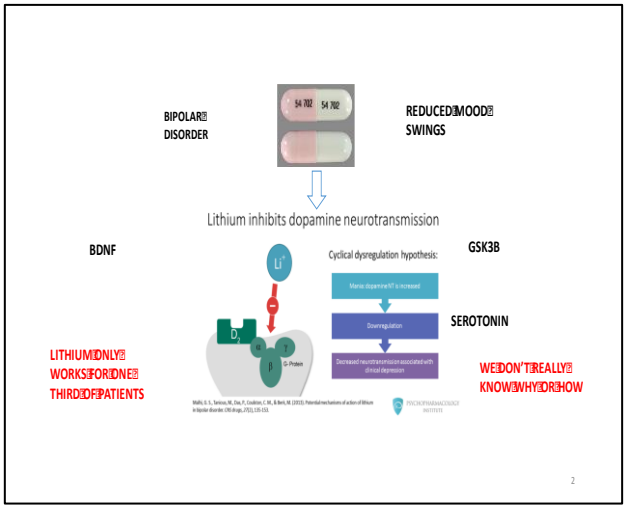
Theory of Change analysis on the one hand is a good way of integrating and synthesising the results of an impact evaluation so as to produce an evidence-based

assessment of whether an intervention has had an effect and in what ways. The added value of Theory of Change analysis is that it can be used to establish 'attribution' (the probability that action X will 'cause' result Y), in situations where experimental approaches cannot be applied, by identifying the 'causal pathways' between an intervention's objectives, its activities, and its expected outcomes and impacts. ToC analysis draws evidence from different sources and stakeholders and triangulates this evidence to identify the factors that are likely to have 'caused' an effect.

How to do it

The key to Theory of Change analysis is identifying and analysing the 'mechanisms' of an intervention - or more precisely the 'Context-Mechanism-Outcome (CMO) configuration'. Mechanisms can be defined as the underlying entities, processes, or structures which operate in particular contexts to generate outcomes (Befani, 2012).[22] A mechanism describes a complex combination of causes, and what their role is in contributing to the project's results. Mechanisms work because of the interaction between two key elements – the two 'R's' – 'Resources' and 'Reasoning'. Resources are the 'contextual assets' provided by a project – for example in Designscapes the Technical and Financial Support instruments provided to pilots. These resources interact with – and subsequently change – the 'reasoning' of the actors involved in the project. The mechanism explains an outcome by describing how an individual's (or group's) 'reasoning' – their thinking, acting or decision-making - is influenced by the 'contextual resources' provided by the project. A good way of understanding a mechanism is to think of it as the 'missing middle' between a project's 'context' (the presenting problem it wants to solve and the resources available to solve it) and the changes to that problem that the project is expected to deliver once it has been completed (the outcomes and impacts).

[22] Befani, B. (2012) 'Models of Causality and Causal Inference', in E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (eds), Broadening the Range of Designs and Methods for Impact Evaluations, DFID Working Paper 38, London: Department for International Development



Take the context of bipolar disorder. A patient presenting with bipolar disorder is given a course of lithium, and the observed outcome of this action is a reduction in the patient's mood swings.

It's tempting to assume that lithium ***causes*** a reduction in mood swings but this is a mis-attribution. Lithium contributes to a complex neurological and neurochemical process that involves the enzyme GSK38, the circadian clock, transcription of brain-derived neurotrophic factor (BDNF) and serotonin. This complex process is the 'mechanism of action'. We don't really know how it works and it only works in around a third of patients.

The example of lithium described above illustrates five key attributes of mechanisms:

- they reveal the 'missing middle' between context and outcomes. To capture the mechanism, you need to open up the 'black box' – a capsule of lithium; an MRI scan of the neurology of someone with bipolar disorder – and see what's inside
- mechanisms are always hidden. You can't see the mechanism of gravity affecting an apple falling from a tree. You can't see how lithium works on neurological systems. If you can see the mechanism, then it's likely to not be a mechanism
- mechanisms – and interventions – are influenced by 'Resources' – a theory of gravitational force; a capsule of lithium
- mechanisms – and interventions – are influenced by, and in turn influence, 'Reasoning' - social and psychological processes by and between actors
- mechanisms – and interventions – do not work all the time. They may only work for some people under some circumstances.

In practice, revealing and working with the mechanisms that define the causal pathways between context and outcomes in an intervention requires three main steps:

- Step 1: Specification - specify the mechanisms that are assumed to contribute to the key impact areas of the intervention and any alternative mechanisms not within the control of the intervention that could also contribute to these key impact areas
- Step 2: Evidence Triangulation – collect the data from the evaluation results that support the proposed mechanisms together with any evidence that

contradicts the proposed mechanisms, as well as evidence that supports the alternative mechanisms

- Step 3: Interpretation - using the results of the evidence triangulation to present an evidence-based assessment of how far the intervention has progressed on its expected 'change journey' and whether the theories and assumptions proposed in the intervention's Theory of Change on what causes change can be supported by the evidence.

The following sections provide a Guideline to take you through these three steps, illustrated by the example of Theory of Change analysis for the 'Designscapes' project.

Step 1: Specification

This step involves specifying the mechanisms that are assumed to contribute to the key impact areas of the intervention and any alternative mechanisms not within the control of the intervention that could also contribute to these key impact areas. The Table below shows the mechanisms and alternative mechanisms developed for the Designscapes key impact areas.

Impact area	Mechanisms	Alternative mechanisms
Competitiveness - design-led innovation improves performance and efficiency in the commercial and public sector, and hence improves competitiveness. The urban context makes an additional contribution.	Designscapes provides a flexible funding instrument to which applicants respond by applying design thinking and tools to generating and developing new and innovative ideas to tackle problems linked to their urban environment. This leads to sustained innovation capacity at urban level and creates economic and social value.	There are lots of mainstream national and European funding streams that focus on finding solutions to pressing social, environmental and economic challenges which combine to spark innovations which solve these problems effectively.
Inclusion - Designscapes supports inclusiveness and reduces inequalities in citizens' access to innovation	By using design thinking and tools, Designscapes funded pilots are enabled to involve a broad spectrum of citizens as	Citizens don't have the knowledge and skills to come up with solutions to the problems they face, and are generally

Impact area	Mechanisms	Alternative mechanisms
	end users of the innovation into the DEI process. This ensures citizens' buy-in into the innovation and encourages them to use the innovation, hence reducing inequalities in access to new products, services or processes.	disengaged from political and other processes. It therefore requires the work of professionals to come up with solutions to complex problems and implement them.
Co-creation - Designscapes supports co-creation, which in turn leads to successful innovation	By using design thinking and tools, end-users' creativity is mobilised and combined with pilot teams' technical and professional expertise. This leads to new solutions or the adaptation of existing ones which meet user needs better and hence get adopted more widely. This generates business and social value.	

Table 8: Mechanisms and alternative mechanisms

Step 2: Evidence triangulation

This step involves collecting the data from the evaluation results that support the proposed mechanisms together with any evidence that contradicts the proposed mechanisms, as well as evidence that supports the alternative mechanisms. For reasons of brevity, this Step is illustrated by the example of the Designscapes project with reference to one impact area and one mechanism only. A full description of the mechanisms can be found in Designscapes D.2.3 – Final Evaluation Report.

Impact Area Co-Creation

	Hypotheses	Evidence
Description:	By using design thinking and tools, end-users' creativity is mobilised and combined with pilot teams' technical and professional expertise. This leads to new solutions or the adaptation of existing ones which meet user needs better and hence get adopted more widely. This generates business and social value.	
Resources:	Call for Designscapes pilot Designscapes funding Design tools and methods used by funded pilots	Text of the Designscapes call emphasises co-creation Successful pilots predominantly use co-creation methods
Immediate outcomes:	Diverse citizens motivated and engaged Knowledge of design thinking, tools and methods Improved knowledge about adapting innovations to different contexts	Application data shows a range of user groups involved Case study data shows that design tools can engage and motivate diverse user groups to work together to create new solutions. Survey data shows pilots gain skills in design tools and feel more capable of adapting the innovation to new contexts
Intermediate outcomes:	Innovations meet user needs and get adopted Pilot teams develop new ways of working Access to funding	Case study data shows citizens engaging with the innovations they have helped create Survey data shows teams develop new ways of working (e.g.collaboration and use of design thinking) Survey data shows pilot teams access funding beyond Designscapes to

		continue to develop their innovations
Long-term impact:	DEI creates sustainable innovation DEI creates social value DEI creates economic value	Case study data shows innovations are created and suggests these may be sustainable Case study data shows citizens involved gain confidence and skills, pilot teams reporting meting social impact objectives. Case study data suggests design-led way of working and co-creation adopted by teams beyond Designscapes and urban stakeholders become more susceptible to this way of working
Assumptions:	Disruptive innovation will be generated because projects will integrate citizens / customers and other stakeholders in innovation processes	Evaluation data shows that innovations were created, though these were not necessarily disruptive.
Alternative mechanism:	Citizens don't have the knowledge and skills to come up with solutions to the problems they face, and are generally disengaged from political and other processes. It therefore requires the work of professionals to come up with solutions to complex problems and implement them.	Alternative mechanism poorly supported by the evidence.

Table 9: Contribution Analysis table co-creation

Step 3: Interpretation

This step involves using the results of the evidence triangulation to present an evidence-based assessment of how far the intervention has progressed on its expected 'change journey' and whether the theories and assumptions proposed in

the intervention's Theory of Change on what causes change can be supported by the evidence.

Contribution analysis

What is it

Contribution Analysis is an approach for assessing causal questions and inferring causality. It offers a step-by-step approach designed to help managers, researchers, and policymakers arrive at conclusions about the contribution their programme has made (or is currently making) to particular outcomes. The essential value of contribution analysis is that it offers an approach designed to reduce uncertainty about the contribution the intervention is making to the observed results through an increased understanding of why the observed results have occurred (or not!) and the roles played by the intervention and other internal and external factors.

http://betterevaluation.org/plan/approach/contribution_analysis

When should it be used

Contribution analysis is used in situations when the evaluation audience and key stakeholders are asking for evidence of the 'counterfactual' in an impact evaluation [23]. Counterfactual impact evaluation involves comparing the outcomes of interest of those who have benefitted from an intervention (the 'treatment group') with those of a group similar in all respects to the treatment group (the 'comparison/control group'), but who have not been exposed to the intervention. The comparison group provides information on what would have happened to the participants in the intervention had they not been exposed to it. As noted above, the preferred approach typically involves an experimental methodology using randomised controlled trials (RCTs) where participants are randomly assigned to the treatment and comparison groups and then common outcomes are compared before and after the intervention for both groups. In evaluations involving complex 'social' interventions it's very difficult to satisfy the conditions required for an experimental approach to be implemented. You can try 'quasi-experimental' methods in this situation, and supplement these with non-experimental methods – for example ethnographic methods – as discussed above. Using a 'mixed method' approach in this way requires a way of putting the evidence together and testing it for its validity and capability of answering the evaluation questions. This is where contribution analysis comes in.

[23] Loi, M and Rodrigues, M (2012) A note on the impact evaluation of public policies: the counterfactual analysis, JRC Scientific and Policy Reports, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC74778/lbna25519enn.pdf

How to do it

Contribution Analysis aims to create a causal chain – or 'contribution story' – that links actions and events to outcomes. The contribution story involves six steps [24]:

1. Set out the attribution problem to be addressed – specifying the outcome or target that is hoped to improve or change, as well as the key evaluation questions to be addressed
2. Develop a theory of change about how the intervention is supposed to work, together with i) the assumptions underpinning the theory ii) the risks to realisation of the intended outcomes and impacts
3. Gather evidence to assess whether the Theory of Change works, and explore and discuss plausible alternative explanations - identifying the most likely alternative explanations and the evidence associated with them
4. Assemble the Contribution Story – explain how and why a result is caused by a particular sequence of events and actions, and why it is reasonable to assume that the actions of the intervention have contributed to the observed outcomes. Specify the weaknesses in the story
5. Gather new evidence on the implementation of the intervention
6. Revise and strengthen the contribution story – using the new evidence gathered and assessed.

The key to a good contribution analysis is to create a story that specifies:

- The Primary Mechanisms - 'underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest'
- The Alternative Mechanisms – 'underlying entities, processes, or structures that are unconnected with the intervention'
- The Influencing Factors – 'contextual conditions that might enable or impede mechanisms'.

To address Criteria 4 and 5, a useful tool to use is the Relevant Explanation Finder (REF) (Lemire et al, 2012) which provides a guiding framework for the identification and examination of influencing factors and alternative explanations. [25] This entails developing 'alternative primary explanatory mechanisms' together with their 'Influencing Factors' to the primary mechanisms identified. Each alternative explanation and influencing factor is described using six dimensions:

- Description of the mechanism
- Type of explanation or factor identified. This covers four types:

[24] Mayne,J. (2012) "Contribution analysis: Coming of age?" *Evaluation*, 18(3), pp.270–280

[25] Lemire, S, Bohni Nielsen, S and Dybdal, L (2102). Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. Evaluation, 18(3) 294–309.

  - Primary explanation (mechanism) – a mechanism identified and purported to be the target intervention mechanism that accounts for and explains the observed outcomes (i.e. the primary explanatory mechanism in the contribution story).
  - Direct rival (mechanism) – a mechanism, different from the target intervention mechanism, that accounts for and explains the observed outcomes (i.e. undermines the contribution story).
  - Commingled rival (mechanism) – other mechanisms, along with the target mechanism, that both contribute to and explain the observed outcomes (i.e. revise the contribution story).
  - Implementation rival (factors) – influencing factors in the implementation process, not substantive intervention mechanisms, that modify the outcomes (i.e. revise the contribution story).
- Explanation level. This also covers four types:
  - The individual level – the capacities of the key actors and stakeholders such as interests, attitudes, capabilities and the credibility of (for instance) professionals or beneficiaries.
  - The interpersonal level – the relationships required to support the intervention, such as lines of communication, management and administrative support, union agreements and professional contracts.
  - The institutional level – the setting in which the intervention is implemented, such as the culture, leadership etc. of the implementing body.
  - The wider (infra-)structural level - political support, the availability of funding resources etc.
- Identifiers – provides descriptions of the possible identifiers that offer proof or disproof of the existence of the various influencing factors and alternative explanations.
  - Degree of influence – this summarizes the degree of influence of both the factors identified and the underlying mechanisms for the observed outcomes. It covers five attributes to be considered in assessing the degree of influence:
  - Certainty : The degree to which the observed outcome pattern matches the one predicted by the factor or mechanism.
  - Robustness : The degree to which the factor or mechanism is identified as a significant contributor across a broad range of data sources and data collection methods.
  - Range : The degree to which the factor or mechanism contributes to a broad range of the outcomes of interest.
  - Prevalence : The degree to which the factor or mechanism contributes to the outcomes of interest across a wide range of implementation environments and target groups (e.g. different implementation sites

and / or types of intervention).

- Theoretical grounding : The factor or mechanism is informed by theory (identifies existing theories of which it is an example) and is cast in specific terms (i.e. it is not vague).

- Implications – this provides conclusions regarding the implications of the factors and mechanisms for the contribution story.

Ultimately, Contribution Analysis aims to infer 'plausible association' between an intervention and a set of relevant outcomes by means of systematic inquiry. To demonstrate this 'plausible association', an intervention 's Theory of Change needs to meet the following five criteria Mayne (2011):

- Plausibility: Is the theory of change plausible?
- Implementation according to plan: Has the program been implemented with high fidelity?
- Evidentiary confirmation of key elements: To what extent are the key elements of the theory of change confirmed by new or existing evidence?
- Identification and examination of other influencing factors: To what extent have other influencing factors been identified and accounted for?
- Disproof of alternative explanations: To what extent have the most relevant alternative explanations been disproved?

PART B: APPLYING THE COMMON IMPACT METHODOLOGY – APPLICATION AREAS

The second part of this Guideline shows how the Common Impact Methodology described in Part A can be applied in practice, using the evaluation of the Designscapes project as an illustrative example. Three application areas of the methodology are covered below. Each application area is linked to a particular objective of the Horizon 2020 SC6-CO-CREATION-2016-2017 Call. Each focuses on a specific evaluation question, and each is illustrated and exemplified by the use of a particular method to evaluate an aspect of the Designscapes project.

Application Area 1: User benefits and business impacts

Call Objective: gather data and metrics concerning the impact of design-related policies and programmes in terms of user benefit and business impact

Example Method: Cost Consequence Analysis (CCA)

Introduction

This application area focuses on assessing the contribution of design-enabled innovation in two areas: user benefits and business impact. The key evaluation question addressed in this example is:

- What user benefits and business impacts are associated with design-related policies and programmes?

The impact evaluation method used to assess user benefits and business impact is Cost Consequence Analysis (CCA). This calculates the economic and social benefits of the funding provided to projects by Designscapes through its financial instrument.

What is Cost Consequence Analysis and when is it used

Cost Consequence Analysis (CCA) is a method used to assess the value returned by an intervention. It's typically used as an alternative to Social Return on Investment (SROI). SROI measures social, environmental and economic outcomes and uses monetary values to represent them [26]. On the surface, calculating an SROI ratio – the amount of return on the financial input invested – is a simple matter of using the following formula:

SROI = (social impact value – initial investment amount) ÷ initial investment amount x 100%

[26] Source: 'A Guide to Social Return on Investment'. SROI Network, 2012

For example an SROI ratio of 5:1 indicates that for every euro invested, an intervention delivers 5 euro in value (defined as economic, social and environmental value). In practice, calculating SROI is quite complex and requires the following steps:

- Establishing the financial proxies needed to calculate the SROI. Proxies are indirect indicators that approximate for a direct indicator for which data is difficult to obtain. For example, there are no common and recognised indicators to measure the financial value attributed to the impact a DEI intervention has on social inclusion. So a proxy measure needs to be found – for example the reduction in average welfare costs expended to support a young person who is NEET (not in Employment, Education or Training)
- Developing an 'impacts map' that shows relationships between inputs, outputs and outcomes – this is normally derived from an intervention's Theory of Change
- Evidencing the inputs, outputs and outcomes and giving them a value
- Establishing impact. This requires calculating 'additionality' - "the extent to which something happens as a result of an intervention that would not have occurred in the absence of the intervention" 27. Calculating additionality requires four sub-calculations: deadweight - a measure of the outcome that would have happened even if the intervention had not taken place; attribution - an assessment of how much the outcome was caused by the contribution of other organisations, interventions or actors; leakage - the value that is lost through, for example, an organisation leaving the intervention; displacement/substitution - the extent to which the benefits claimed by participants are at the expense of others outside the intervention; drop-off - the deterioration in the effects of an outcome over time.
- Calculating the SROI. All the benefits of the intervention are aggregated. These are set against costs and 'negative' values and the result compared with the financial investment.

In the DEI field – because many initiatives are innovative experiments that 'think out of the box' – it's often impractical to carry out an SROI for reasons like the lack of relevant proxy measures; lack of quantitative data on outcomes and the absence of baselines to measure attribution. CCA is used in these situations. Like SROI, CCA considers a broader range of outcomes and measurements beyond financial ones, including 'humanistic' measures, for example in the health field indirect medical costs such as changes in productivity – like earnings lost because of illness – the costs of pain, anxiety, loss of self-esteem, and inability to carry out normal activities. The big difference between CCA and SROI is that CCA "does not attempt to summarise

27 'Additionality Guide - third edition'. English Partnerships, 2008

outcomes in a single measure (such as the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out". 28 Another difference is that CCA takes a broader view of the 'counter-factual' - what is likely to happen in the absence of the intervention – than SROI, which uses a monetised calculation of additionality.

CCA is generally used to compare two scenarios – the current status quo and an alternative scenario represented by the intervention. It's easier to apply comparative CCA in cases where the boundaries and parameters of the status quo and the intervention are clearly defined. For example, in the DEI field, the simplest application context would be at the organisational or SME scale – for example comparing the situation before an organisation introduced a design-enabled change to its systems and practices with the situation afterwards. At the macro scale the conditions required for CCA become more difficult to satisfy – because, for example, it's hard to collect comparative costs across entire industries or nations.

How to do CCA

CCA analysis involves five stages, summarised in the Table below.

Stage	What this involves
1. Mapping	Produce an Impacts Map showing the expected impacts of the changes realised by the intervention
2. Measurable financial consequences	Select and quantify the impacts and outcomes that have measurable financial consequences attached
3. Non-financial consequences	Select and evaluate the impacts and outcomes that have measurable non-financial consequences attached
4. Non-recurrent costs	Identify and quantify non-recurrent costs
5. Analysis	Compare cost consequences of alternatives and review results

Table 10: Stages in CAA

Applied to Designscapes, CCA was used to calculate the economic and social benefits of the funding provided to projects by Designscapes through its financial instrument, which supported over 100 initiatives in European cities to develop, pilot

28 https://www.nice.org.uk/Glossary?letter=C

and potentially scale design-enabled innovations addressing pressing social, environmental and economic challenges. The financial instrument was delivered in three Calls:

- Call 1 – supported 50 projects with a 'sub-grant' of € 5.000 each to produce a feasibility study for a DEI project, defined as 'a description of a proposed project or initiative, outlining its main (expected, or desired) characteristics, a time plan for involved activities, a cost and value creation analysis, and an assessment of its practicality'.
- Call 2 - supported 41 projects with a 'sub-grant' of € 25.000 each to deliver a DEI prototype, defined as 'an experimental release of a new product, service, process or other innovative solution, built according to a predefined guideline (including a feasibility study) and tested in a laboratory environment and/or in real life conditions, with or without the participation of its prospective end users.'
- Call 3 - supported 10 projects with a 'sub-grant' of € 25.000 each to produce a 'scalability proof' for a DEI project, defined as 'a special project or initiative, materializing the successful replication, and/or transfer, and/or reuse, and/or diffusion of an existing prototype, in one or more additional contexts than those where it was originally conceived, implemented or tested.'

The main dataset for the Designscapes CCA is drawn from a survey of funding applications under Calls 2 and 3. Just over 100 survey responses were collected from actors involved in these applications, broken down as follows:

- 39 respondents from projects funded under Call 2 (prototypes)
- 19 respondents from unsuccessful applications – i.e. projects that were not funded – under Call 2 (prototypes)
- 13 respondents from projects funded under Call 3 (scalability proofs)
- 30 respondents from unsuccessful applications – i.e. projects that were not funded – under Call 3 (scalability proofs).

CCA was used in 'comparative' mode in the Designscapes application to compare these two sets of projects – i.e. successful v unsuccessful. This enables a counterfactual assessment to be established.

The Impacts Map produced for the CCA (Stage 1) draws on the Designscapes Theory of Change with a particular focus on two impact areas, as specified in Objective a) of the Horizon 2020 SC6-CO-CREATION-2016-2017 Call: user benefits and business impacts. The Impacts Map is shown in Table ?.

Impact area	Mapping outcomes
User benefit	Involved citizens in DEI Contribution to societal challenges (social exclusion, climate change,

Impact area	Mapping outcomes
	economic, crisis of democracy, crisis of values, quality of life) and financial value of contribution Developed design-thinking skills
Business Impact	Accessed new funding sources Generated additional funding Stimulated Business Network-building Expanded professional networks Developed links with potential customers Developed user base Developed new ways of working Financial value of participating in Designscapes

Table 11: CCA mapping outcomes against Designscapes impact areas

Stage 4 - Identify and quantify non-recurrent costs – was not covered in this CCA because the units of analysis are prototypes and scalability proofs which are not yet at a stage in their life cycle in which they have been established and are incurring non-recurrent costs.

Results of Cost Consequence Analysis of the Designscapes project

The results of the Cost Consequence Analysis are shown in Table ?. Table ? compares two Designscapes scenarios – a 'successfuly funded' scenario against 'not funded' scenario. The two scenarios are compared against the two key impact areas: user benefits and business impact.

Impact Area: User benefits	
Successful applicants	**Unsuccessful applicants**
Call 3 projects estimate Designscapes contribution to addressing societal challenges on average €110,000	No data on estimated value of contribution to addressing societal challenges

Call 2 projects estimate Designscapes contribution to addressing societal challenges on average €1.12m 60% of projects reported Designscapes funding had supported them to meet their societal challenges targets to a great extent or completely The biggest impact was on social inclusion (85% reported addressed to a great extent or completely), quality of life (77%), crisis of democracy and crisis of values (62% respectively). Significantly less impact was reported for climate change (31%) and economic impact (38%) Involvement of citizens – 69% Call 3 projects report involving citizens, at average of 230 per project . 90-% Call 2 projects report involving citizens, at average of 89 per project . Developing skills in design thinking – 77% Call 3 projects report developing new design thinking skills completely or to a large extent 88% Call 2 projects report feeling extremely or very confident on their ability to use design methods or tools	Only 18% Call 3 applicants reported their projects had addressed the six societal challenges targeted by Designscapes Impact on societal projects reported by Call 3 applicants low across the six challenge areas - addressed to a great extent or completely: quality of life (23%), crisis of democracy and crisis of values (17% and 13%respectively); climate change (20%) and economic impact (13%) No data on numbers of citizens involved. 23% Call 3 applicants report involving citizens in their projects. 84% Call 2 applicants report involving citizens in their projects. 33% of Call 3 funded projects report developing new design thinking skills completely or to a large extent 89% Call 2 projects report feeling extremely or very confident on their ability to use design methods or tools
Impact Area: Business Impact	
Successful applicants	**Unsuccessful applicants**
Call 3 applicants average €39,500 estimated financial value of participating. Leverage: 1.58. Estimated return on project: €80,000 Call 2 applicants average €33,708 estimated financial value of participating. Leverage: 1.34	Call 3 applicants average €21,761 estimated financial value of participating – 81% less than Call 3 successful applicants Estimated return on project: €76,000

23% Call 3 applicants report generating additional funding – on average €16,170 per applicant 21% Call 2 applicants report generating additional funding – on average €74,428 per applicant Call 3 and Call 2 projects report additional business impact benefits in collaborating with partners, identifying new customers, and developing new ways of working. Call 2 projects report additional impact on expanding user base.	17% Call 3 applicants report generating additional funding – on average €26,600 per applicant 26% Call 2 applicants report generating additional funding – on average €20,714 per applicant Non-successful applicants report significantly lower levels of business impact benefit in these areas, and no significant benefit in other areas

Table 12: CCA comparing successful and unsuccessful Designscapes applicants

The key conclusions from the CCA with regard to ***business impacts*** are:

- The analysis shows that Designscapes funding acted as a financial stimulus for funded projects. Both Call 3 and Call 2 successful applicants reported an average estimated financial value of participating in Designscapes of €39,500 and €33,708 respectively - at a leverage rate of 1.58 and 1.34.
- Similarly, both Call 3 and Call 2 successful applicants reported generating additional funding – an average €16,170 – at a leverage rate of 0.65 - for Call3 applicants – and an average of €74,428 – at a much higher leverage rate of 2.9 - for Call 2 applicants.
- The analysis suggests that participation in Designscapes generated greater financial reward than non-participation. Successful applicants estimated the financial value of participating at a level 81% higher than unsuccessful applicants. The average estimated financial return on the project was €80,000 for successful Call 3 applicants compared with €76,000 for unsuccessful applicants.
- This latter finding shows, however, that unsuccessful applicants were still able to generate financial value for their projects – as a result of looking beyond Designscapes. Although a smaller proportion of unsuccessful Call 3 applicants were able to secure additional funding than successful applicants, they secured a higher rate of funding on average than successful Call 3 applicants. This situation was reversed for Call 2 applicants, with successful applicants securing a much higher level of funding on average than unsuccessful applicants.
- Participation in Designscapes also appears to have been associated with other business benefits in addition to a purely financial return, in particular contributing to collaborating with professional partners, identifying new

customers, and developing new ways of working. This pattern was reported for both Call 3 and Call 2 funded projects – although Call 2 projects reported a lower level of benefit with regard to identifying new customers. Call 3 projects reported less benefit associated with expanding their user base – although this aspect of business impact was rated significantly higher by Call 2 projects.

- Across the board, unsuccessful Call 3 applicants reported much lower levels of additional business impact benefits than the successful Call 3 projects. No data were available on these measures for unsuccessful Call 2 applicants.
- On balance , therefore, the Cost Consequence Analysis suggests that Designscapes generated business impacts that would not have been produced if the project had not been implemented – although on a relatively modest scale.

The key conclusions from the CCA with regard to ***user benefits*** are:

- Projects funded under Call 3 report Designscapes supported them to make a significant contribution to addressing societal challenges, with a value estimated at €110,000 on average per project. Across the six societal challenges targeted by Designscapes, 60% of funded projects reported the funding had helped them achieve their expected impacts to a large extent or completely. For projects funded under Call 2, the reported impact of Designscapes funding on societal challenges was much greater, estimated at €1.12m on average per project.
- However, the Designscapes contribution to societal impacts was not uniform. The biggest impact reported by funded Call 3 projects on societal challenges was on social inclusion (85% of funded projects reported meeting their targets to a great extent or completely), quality of life (77%), crisis of democracy and crisis of values (62% respectively). Significantly less impact was reported for climate change (31%) and economic impact (38%).
- In contrast of the unsuccessful Call 3 applicants only 18% reported they had secured funding that aimed to address the societal challenges covered by Designscapes.
- Another key impact area addressed by Designscapes was the involvement of citizens in design-enabled co-creation. The CAA showed that Desigsncapes funding appears to have made a significant contribution to supporting citizen involvement. 69% of Call 3 funded projects and 90% of Call 2 funded projects reported involving citizens in co-design, compared with 23% of unsuccessful Call 3 applicants – although the proportion of unsuccessful Call 2 applicants involving citizens in co-design – at 84% - was almost as high as for successful Call 2 projects.
- This finding is linked to developing skills in using design thinking to support innovation. The CAA showed that 77% of Call 3 funded projects report developing new design thinking skills completely or to a large extent, compared with 33% of unsuccessful Call 3 applicants. However, for Call 2, the impact on design thinking is much less pronounced. When asked to rate their

confidence on their ability to use design methods or tools, 88% of funded Call 2 projects on average reported feeling extremely or very confident, compared with 89% of unsuccessful Round 2 applicants.

- On balance , therefore, the Cost Consequence Analysis suggests that Designscapes generated user benefits that would not have been produced if the project had not been implemented – although this impact was not uniform across the dimensions assessed in the analysis.
- In addition, the CCA findings suggest that Desigsncapes had a greater impact on user benefits at the 'scalability proof' level than at the 'prototype' level. This is likely to be due to the greater maturity of projects funded under Call 3, which meant they were further along their 'change journey' than Call 2 projects, and thereby further along the road to achieving impacts.

Limitations of the analysis

The analysis presented above needs to be accompanied by a 'health warning'. Although every effort has been made, through the methodology adopted, to be as rigorous as possible in producing the calculations for the CCA, it should be acknowledged that the CCA method reflects a degree of subjectivity and interpretation (as is acknowledged in the literature on it use). The CCA analysis relies heavily on the data derived from a Survey of projects funded through the Designscapes funding instrument – compared with a survey of applicants who applied for funding but were not successful. The survey population is relatively small and this places some limitations on the generalizability of the survey results (although the sample size is large enough for statistical inferences to be made). The analysis also mixes different units of analysis – prototype applications under Call 2 and scalability proof projects funded under Call 3. Since the scalability projects are at a different stage in development, they are more likely to have achieved measurable outcomes than the Call 2 applicants. Another issue is that the monetised data analysed in the CAA – in particular the financial contribution attributed to participation and the estimate of the financail impact associated with addressing societal challenges – is based on self-reported estimates. In addition, the analysis would have benefited from additional sources of comparative data, in order to improve triangulation. For example, case studies were carried out of funded projects but not for unsuccessful applications.

Application Area 2: The Relationship between design and innovation

Call Objective: develop a transferable methodology to evaluate the effectiveness of design in the innovation process

Example Method: Regression analysis

Introduction

This application area focuses on assessing the relationship between design and innovation. The key evaluation question addressed in this example is:

- What contribution does design thinking make to innovation?

The impact evaluation method used to explore the relationship between design and innovation is Regression Analysis. This method is used to predict the likely effects of the application of design thinking methods and tools to increase innovation.

What is Regression Analysis and when is it used

Regression analysis is a statistical analysis method that fulfils a number of functions of interest to impact evaluation including:

- Providing an explanation of the causal relationships between variables and how they lead to observed outcomes
- Providing a measure of the relative strength of the contribution each variable of interest makes to the observed outcomes
- Making predictions about the likely effects on outcomes of interest if the values of the contributing variables are changed
- Producing simulation models based on 'what-if' scenarios – for example if you increase the level of investment design-enabled innovation, what is the likely effect on public sector competitiveness?

How to do Regression Analysis

In regression analysis we try to estimate the relationship between the target variable (the dependent variables) and the independent variables (the predictor variables). Regression analysis fits a function on the available data. If the function is a reasonable fit it can be used to predict the outcome (the dependent variable) in the future.

There are several **types of regressions**, these partly depend on the type of data available. The most simple regression is the linear regression. Linear regression uses a linear function to predict the dependent variable:

$$Y = \beta o + \beta 1X + \in$$

Where:

- Y = The dependent variable, we want to predict
- X = The independent variable, we use to make a prediction
- βo = The intercept coefficient; the prediction when X = 0.
- β1 = slope coefficient; the change in Y when X changes by 1 unit.
- ϵ = The error term, i.e. the difference between actual and predicted values.

Other types of regression include the logistic regression. Here, the dependent variable only has two possible values. In the case of Designscapes an example is the success of applications. We could have developed a regression model to predict the likelihood of success based on predictor variables.

As with all statistical tests, a number of **assumptions** need to be met in order for a regression analysis to be carried out. For a linear regression these assumptions are:

1. The independent and dependent variable need to be metric (i.e., measured at the continuous level).
2. The relationship between the dependent and the independent variable needs to be linear.
3. The dependent variable and the error terms must be normally distributed.
4. The error terms need to contain constant variance. If this is not met, it leads to heteroskedestacity instead of homoscedasticity.
5. The error terms must be uncorrelated.

There is no common agreement on the required **sample size** for regression analysis. Guidance sets out that the data should consists of at least 100 cases. The more complex the regression the more cases are needed. Roughly 10 cases for every independent variable should be available. If the sample size is too small there is a risk of overfitting the model and hence resulting in an unreliable model.

Regression analysis involves several **steps**. These are summarised into the table below.

Stage	What this involves
1. The problem	Identification of the problem and need for regression analysis; design of research instruments and data collection.
2. The data	Exploration of data, including identification of potential issues (e.g., missing data) and exploration of relationships between variables
3. The variables	Define the independent variable(s) and the dependent variable

Stage	What this involves
4. The model	Analyse models and select model with the best fit
5. Validation	Test assumptions; use best model to predict the outcome based on the predictor variables included in the model

Table 13: Steps in regression analysis

Results of Regression Analysis of the Designscapes project

Applied to Designscapes, regression analysis was used to predict the application of design thinking methods and tools to increase innovation. The linear regression was extended to the multiple regression, i.e. inclusion of more than one independent variable.

Data for the dependent variable was based on the post-surveys that all three calls completed towards the end or after their funding period. The application data was used for the definition of the predictor variables.

Dependent variable "Innovation" was defined as a combination of:

- Increased design knowledge as measured by an item part of the post survey "Please indicate the extent to which you agree or disagree that since applying for funding from Designscapes you have: Deepened your knowledge of design methods and tools".[29]
- Using new ways of working as measured by an item part of the post survey "Please indicate the extent to which you agree or disagree that since applying for funding from Designscapes you have: Tried out new ways of design led working".[30]

A range of **independent variables** were included in the initial model as there were assumed to have a relationship with the dependent variable. All independent variables are based on the application data, hence on responses from applicants before project delivery. The only exception is 'capacity building', this was reported as part of the post survey.
For the Designscapes regression analysis, these independent variables were:

Variable	Measured by

[29] The items and rating scales varied slightly across calls (e.g., in the second call the item was phrased "increased knowledge of using design thinking to address social, environmental, or economic problems" and in the third call "Develop skills in using design thinking and tools for scaling").

[30] Again, items and rating scales utilized varied across call (e.g., the phrasing of the item in the second call was "Stimulating new ways of working" and in the third call "Develop new ways of working").

Team size	Number of team members
Profit orientation	'For profit' or 'not for profit'
Urban context - Institutional capacity	Average across five items (e.g., 'Local facilities and services')
Urban context - Cultural vibe	Average across three items (e.g., 'The cultural debate in our context')
Urban context - Entrepreneurial culture	Average across three items (e.g., 'Support from business associations or networks')
Urban context - Social activisms and integration	Average across two items (e.g., 'An active and involved local community')
Urban context - Environmental awareness	Average across three items (e.g., 'A diffuse sensitivity towards environmental issues')
Designscapes call	'Call 1' or 'Call 2' or 'Call 3'
Gender ratio	Ratio of female tembers
Main target	'Targeted at organisations/regional' or 'national/global target'
Design tools - Participatory design	'Planning to apply' or 'not planning to apply'
Design tools - Design methods	'Planning to apply' or 'not planning to apply'
Design tools - Prototyping methods	'Planning to apply' or 'not planning to apply'
Design tools - Participatory process	'Planning to apply' or 'not planning to apply'
Design tools - Usability evaluation	'Planning to apply' or 'not planning to apply'
User involvement	'Planned to involve socially excluded people/groups'[31] or 'did not plan to involve socially excluded people/groups'.
Capacity building	'Participated' or 'not participated'

Table 14: independent variables used for the Designscapes regression analysis

Using the backwards regression approach that starts with all variables and eliminates the variables that explain the least amount of variance in the dependent variables, results in the predictors in the table below. The model predicting innovation is statistically significant, $R^2 = .34$, $F(4, 64) = 9.917$, $p < .001$. Adjusted R^2 indicates that 34% of the variance in the outcome variable is explained by the model, while the majority of the variance cannot be explained.

[31] Socially excluded groups were defined as 'People with disabilities/impairments', 'Elderly people', '(Second-generation) migrants', and '(Mental) health patients or their carers'

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	**Std. Error**	**Beta**		
(Constant)	2.022	0.381		5.307	0
Urban context: Cultural vibe	0.372	0.095	0.415	3.935	0
Call 2 project	0.603	0.152	0.419	3.969	0
Prototyping methods	-0.433	0.168	-0.279	-2.574	0.012
Project planned to involve socially excluded group	-0.293	0.162	-0.186	-1.809	0.075

Table 15: Designscapes regression Statistics

The coefficients from the table above result in the following regression equation:

$$Y(\text{Innovation}) = 2.022 + .373 * X_{\text{cultural vibe}} + .603 * X_{\text{call 2}} + (-.433) * X_{\text{prototyping methods}} + (-.293) * X_{\text{sociallyexcluded}}$$

Limitations of the analysis

The regression analysis comes with several health warnings and results should be treated with caution and only for exemplary purposes. One of the main limitations of the regression analysis is the small sample size of 70. Therefore, the recommendation of at least 100 observations and 10 per added independent variable is not met. The small sample size also meant that we were unable to divide the data into a train and test dataset. The second main limitation relates to the available data. We were only able to use variables that were available across all three calls. Most predictor variables were based on the application data, but we had limited information about the project delivery, hence the model is based on what projects planned to deliver rather than what they reported to have delivered. All data is also self-reported data. For the dependent data items and rating scales were not the same across calls. We were also unable to include some variables into the regression analysis as assumptions were not met, this was for example the case for 'expertise level of the project team'.

Application Area 3: Value-creating networks and their contribution to efficiency and competition

Call Objective: develop a methodology on how actors in different sectors can better connect with design-enabled innovation to increase efficiency and competitiveness in their respective sectors

Example Method: Ecosystems mapping

Introduction

This application area focuses on the ways in which value-creating networks are stimulated and supported through design-enabled innovation. The key evaluation question addressed in this example is:

- How does DEI support value-creating networks and in what ways do they increase efficiency and competitiveness?

In this example we present the use of ecosystems mapping as a method to assess how the Designscapes Financial Instrument and the funding it provided to support design-enabled prototype and scalability proof projects contributed to creating value-driven networks and the impact these had on efficiency and competitiveness.

What is ecosystems mapping and when is it used

Design research has found that design policy should be based on an analysis of the Design Ecosystem, that is, the constellation of factors which influence the success of design driven innovation. Ecosystems mapping is an established methodology in product and UX design and can be useful in identifying components of a design ecosystem. The ecosystem map is a synthetic representation which attempts to capture the key roles that have an influence on the user, organisation and service environment.[32] An ecosystem map revolves around an identified unit, such as a team or product and is built by first displaying all the entities, and then connecting them based on the type of value they exchange. It can be used to uncover existing gaps and identify valuable opportunities for synergies. Ecosystem mapping is a useful tool that can be used to visualize a product and the systems, processes, and institutions that surround it.

The ecosystem map methodology has been adapted to fit the Designscapes intervention as the ecosystem that developed is purposefully built around the notion of design-enabled innovation. Given this, the map is structured according to findings from the research of Whicher (2017). Whicher found that there are nine components of a design ecosystem, which are the various systemic aspects that are needed for a

[32] Ecosystem Map. Service Design Tools. Available at: https://servicedesigntools.org/tools/ecosystem-map Last accessed 22/10/21.

healthy design ecosystem. These components are: (1) users, (2) support, (3) promotion, (4) actors, (5) designers, (6) education, (7) research, (8), funding and (9) policy. This model is shown in **Error! Reference source not found.** below.

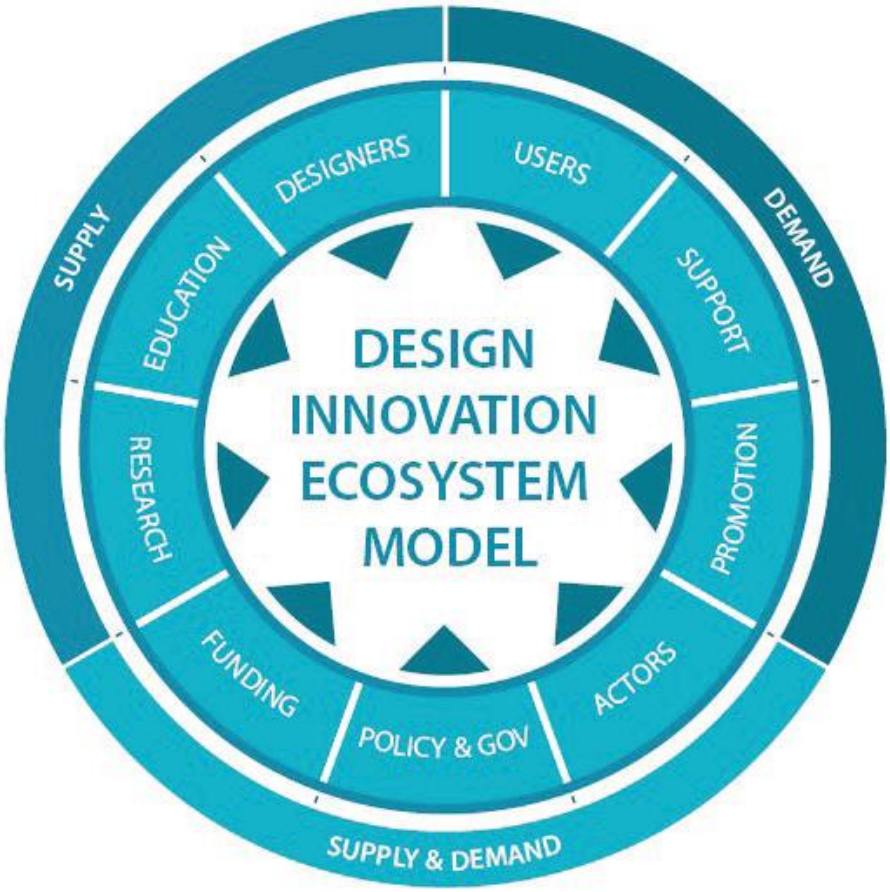


Figure 10: Whitcher's (2017) Design ecosystem model

An ecosystem map based on this model is useful for those pursuing design enabled innovation as it can cluster the roles and contributions that various stakeholders have to a DEI, to identify the gaps and strengths of the network that supports the intervention. An ecosystems map can also represent how close and intense a stakeholder is to the team or product by proximity to the centre.

How to do ecosystems mapping

Mapping an ecosystem can be completed in five steps.

First, identify all players and entities involved in the service supply and delivery. This may require searching in depth for those operating behind the scenes, as well as the

ones not directly involved with the service but that have nevertheless an impact on it.

Second, place the user at the center of the worksheet, then position the other players in the space around. Try to put the ones that have more importance or those with more intense relationships for the service closer to the centre. You may also wish to trace connections between the user and other players and then among the latter. You can also use different types of line (appearance, colour, stroke) to represent different kinds of relationship such as information exchange, money exchange, etc.

Third, when they are all on the map, start arranging the connections by the nine design ecosystems factors identified by Whicher (2017). These are users of the service, support, promotion, designers, education supporting design, research into the product, actors who link to the product, funding for DEI and policy.

Fourth, take a step backwards and observe the map to check whether anything is missing. If you are not just by yourself, discuss within the team. At this point it may be useful to add erasable hand-written notes or use post-its, in order to keep note of emerging observations and discussions.

Fifth, drawing on the new additions, highlight pain points or gaps in the map. This activity helps in understanding which aspects need to be improved or further explored, and to share these insights with other team members or stakeholders. Keep visualising any new detail noticed on the map, for example using further colors or dividing the players in further sub-sets.

Example from Designscapes Impact Evaluation

Figure 11 gives an example from the Designscapes impact evaluation of a design ecosystem map. This map shows the main relationships that existed and were developed during the third phase of the Designscapes funding call for a project called Agroplaza. The map shows that Designscapes was effective in funding projects that were able to draw on resources from across most components of the design ecosystem, and were themselves important in filling gaps through training and funding. This figure was built using Adobe Illustrator but other design tools such as PowerPoint can also be used.

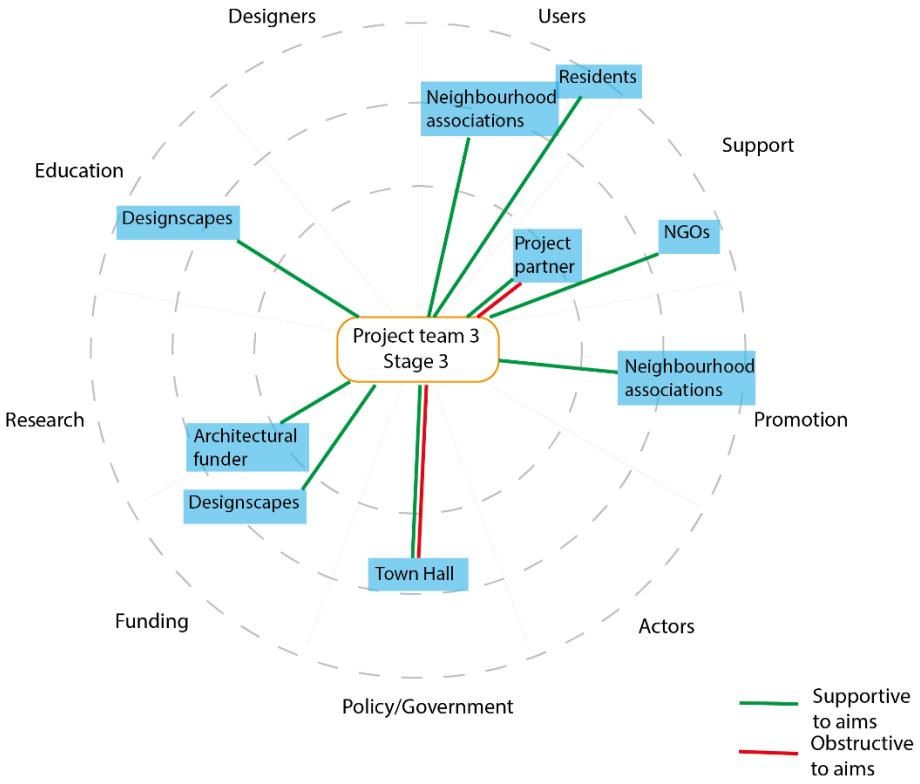


Figure 11: Illustration of ecosystem mapping

References

Astbury B and Leeuw F (2010) Unpacking black boxes: mechanisms and theory building in evaluation. American Journal of Evaluation 31(3): 363–81.

Befani, B, C Barnett and E Stern (2014). Rethinking Impact Evaluation for Development. IDS. Volume 45, Issue 6, pages 17–36, November 2014

Befani, B. (2012) 'Models of Causality and Causal Inference', in E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (eds), Broadening the Range of Designs and Methods for Impact Evaluations, DFID Working Paper 38, London: Department for International Development

Campbell, D.T. and J.C. Stanley, Experimental and Quasi-experimental designs for Research. Chicago, Rand-McNally, 1973.

Chen, H-T (1990). Theory-driven evaluations. London: Sage Publications

Ecosystem Map. Service Design Tools. Available at: https://servicedesigntools.org/tools/ecosystem-map Last accessed 22/10/21.

Guba E and Lincoln Y (1989). Fourth Generation Evaluation. London: Sage Publications

Gobble, M (2014) Design Thinking, Research-Technology Management, 57:3, 59-62

Halfpenny, P (1982) *Positivism and sociology, explaining social life*, London: George Allen and Unwin

Kuhn, T S (1962); *the structure of scientific revolutions.* Chicago: University of Chicago Press

Lemire, S, Bohni Nielsen, S and Dybdal, L (2102). Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. Evaluation, 18(3) 294–309.

Loi, M and Rodrigues, M (2012) A note on the impact evaluation of public policies: the counterfactual analysis, JRC Scientific and Policy Reports, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC74778/lbna25519enn.pdf

Mayne,J. (2012) "Contribution analysis: Coming of age?" Evaluation, 18(3), pp.270–280

O' Donoghue, T. And Punch K. (2003). Qualitative Educational Research In Action: Doing And Reflecting. London: Routledge

Patton M Q (1999). Enhancing the quality and credibility of qualitative analysis. Health Serv Res., 34(5 Pt 2):1189-208.

Pawson, R. and N. Tilley (1997) Realistic Evaluation. London: SAGE

Rittel, H W and Webber, M M (1973). Dilemmas in a General Theory of Planning. Policy Sciences, Vol. 4, No. 2

Stromsdorfer, E.W. (1987).Evaluating CETA: Advances In Assessing Net Program Impact. Evaluation Review, 11:4.

Websites:

https://openknowledge.worldbank.org/handle/10986/25030

https://www.betterevaluation.org/sites/default/files/EA_PM%26E_toolkit_module_2_objectives%26indicators_for_publication.pdf

https://www.ideo.com/post/design-kit