Approaches, ToR and methods for impact evaluation

Q&A
September 2016

European Regional Development Fund

Disclaimer: You are permitted to print or download extracts from this material for your personal use. This material can be used for public use, provided the source is acknowledged and the publisher is given a prior notice. None of this material may be used for commercial purposes. The information and views set out in this interact document reflect Interact's opinions. Responsibility for the information and views set out in this document lies entirely with Interact.
Publisher Interact Programme **Month/Year** September 2016 **Publication Leader** Daniela Minichberger
Contributors Ivana Lazic, Aleksandra Rapp, Philipp Schwartz **Designer** Rosebud Design GmbH

www.interact-eu.net

Table of Content

Introduction

This publication is a practical document which aims at supporting the Interreg programmes in their planning and implementation of impact evaluations in the 2014-2020 programming period.

The document summarises:

- The most important information related to different approaches, methods and techniques for impact evaluation as well as information on drafting Terms of References for impact evaluations for Interreg programmes in the period 2014-2020,
- A set of questions, which were discussed in the context of an evaluation seminar run by Interact Impact Evaluation: Methods and ToR, Amsterdam, 21-22 June 2016,
- The feedback and tips, which were shared by an evaluation expert (Simon Pringle, SQW Ltd), and the Interreg programmes during the seminar.

More information

This publication is considered a working paper which Interact will continuously update. If you would like to comment or contribute to the document, please feel free to contact Daniela Minichberger: daniela.minichberger@interact-eu.net

1. Background of Evaluation, Theory Concept, Logic Model

Increasingly, evaluation is seen by the European Commission as a core discipline in supporting the design, delivery, and performance of programmes. Referring to past good evaluation work helps us understand where and how risk might be mitigated in the design of our new programmes, and ex-ante evaluation provides us with a good understanding of where we start from with our new initiatives. Evaluations during the implementation phase can help us understand whether programmes are doing what they are supposed to and at the end of a programmes' life, ex-post evaluation helps us understand the changes and added value we have brought – *'have we done the right things, and have we been doing things right?'*

Compared to the former programming period there have been some changes in the understanding and organisation of evaluation. *"The most important one is the emphasis on a clearer articulation of the policy objectives. This is key to implement a result oriented policy and move away from an excessive focus on the absorption of funding."*[1]

In order to show this result-orientation **two key concepts now underpin most evaluation approaches**:

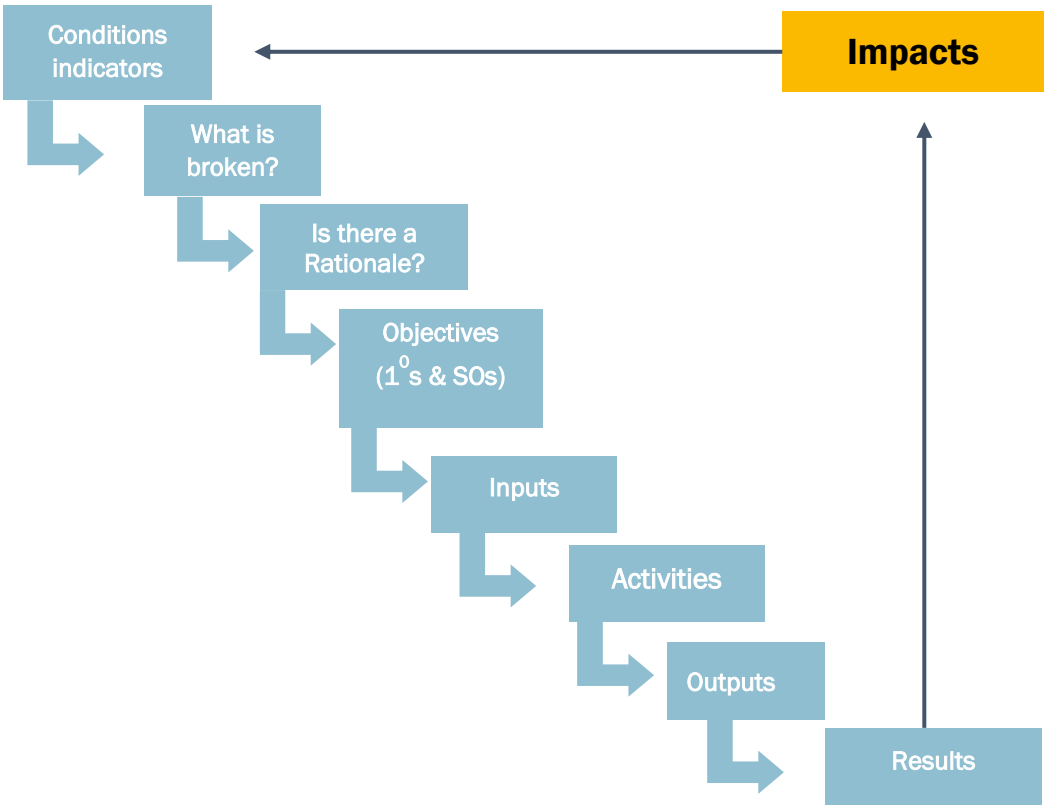
- **The 'logic chain', or logic model**. Typically assembled with eight component parts, as set out in the figure below, the programme's logic chain describes what we expect the intervention to be doing. Starting with the Conditions we are facing and the Problems (also known as market and other failures) in play, the logic chain describes factually what each component involves, set within the context of the other components. Logic chains need to be crisp in content, simple in logic, and clear in their sequential flow. A good logic chain will be accessible and understandable to all programme personnel and partners.

- A **'theory of change'** takes the simple logic chain and moves it from being a descriptive device to provide an explanatory and predictive statement of how, and why the component will perform, deliver, and interface in achieving the programme's success. A good theory of change takes the components of the simple logic model (see example below and Annex 3)[2] and then around these wraps the assumptions and hypotheses about how the programme will work. The assumptions and hypothesis that are used in constructing the theory of change are then routinely tested and explored through evaluative activity.

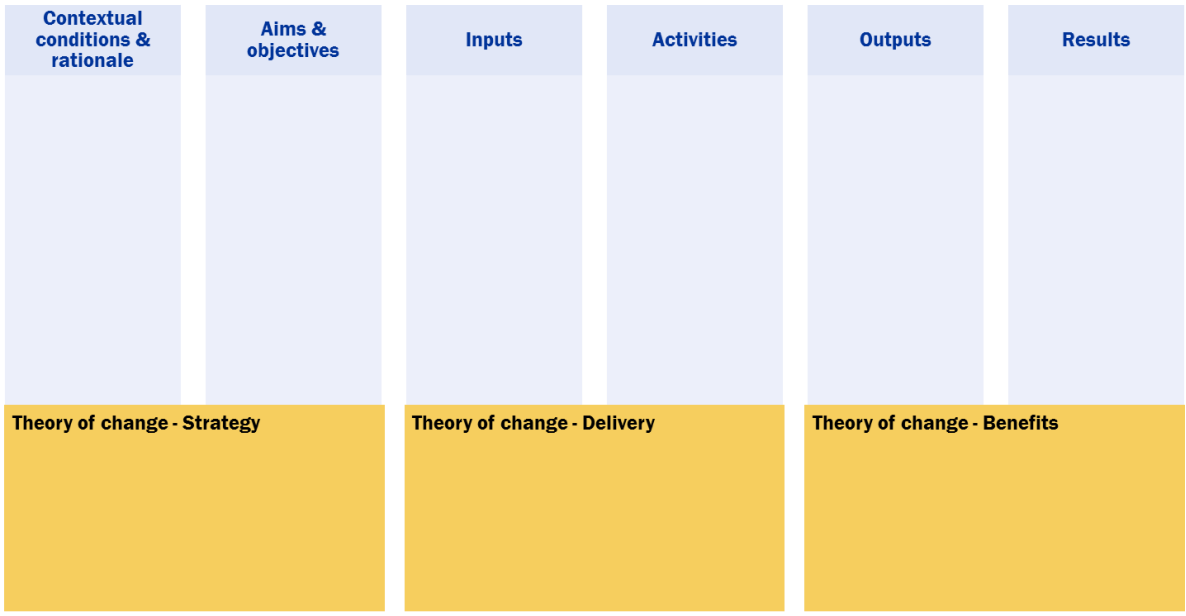
All well-managed programs will have current and up-to-date logic chains and theories of change at all points in their life-cycle —be this at design, approval, implementation and delivery, and close. Both the logic chain and the theory of change are vital building blocks for any evaluation work.

[1] The Programming Period 2014-2020. Guidance document on Monitoring and Evaluation. Cohesion Fund and European Regional Development Fund, Concepts and Recommendations. March 2014, p.2
[2] There are simple and more complicatd ways of displaying the logic model; see Annex 3 for different examples

Logic model - the 'Building Blocks'



Example: simple logic model



Q: Referring to the different logic models - such as very complicated and very simple- would you say that the whole logic of the programme should be reflected in a detailed way? Or is it maybe the challenge to actually make the logic order simple not to get it too complicated?

A (Simon Pringle, evaluation expert): I'm a real believer in keeping things simple and avoiding unnecessary complexity. My maxim as a programme designer and evaluator is KISS – keep it simple and straightforward — so that all those stakeholders involved with the programme - you as Programme Managers, other funders, delivering partners, and even potential applicants — understand what the programme in question is trying to do. So, the starting point for most programmes should be a simple and clear logic chain, out of which a relatively simple theory of change can be developed. Developing logic chains and theory of change is quite a demanding task — you need to think quite hard about what you're doing — so we are not going to produce absolutely first-class products from the outset. Take it simple, and make it straightforward.

Then, as your familiarity with logic chain and theory of change thinking builds, and your operational understanding of the programme in question develops as experience builds, think about reiterating your logic chain and theory of change to be a little bit deeper and more detailed. And continue to think about developing and deepening these devices as your programme continues to mature, and still further experience becomes available to you.

As with all things in life, avoid the temptation 'to run before you can walk'. Start with simple and straightforward logic chains and theories of change, and over time then build these out with further detail and sophistication.

Q: A very crucial point of the logic chain is the rationale. And sometimes the rationale is not very well defined. If you only have a weak rationale, a hypothesis, and you don't know if it's right or wrong, are then are all the others steps weak?

A (Simon Pringle, evaluation expert): There are two foundational blocks in the logic chain: the rationale for intervention (i.e., what is the case for us to intervene through Interreg) and directly related to this the objectives to be secured by the programme. You need to think hard about both elements and make sure that these are sound, convincing, specific, and realistic in your logic chains and subsequent theories of change.

With a sound rationale and set of objectives in place, then the following parts of the logic chain are defined with relative ease, and a strong and convincing programme story is the likely result. By contrast, a poorly defined rationale and set of objectives — essentially meaning we don't really quite understand the 'what', 'why', and 'how' of what we're trying to do — may mean, no matter how elegant our activities, outputs, and impacts, that our programme is ultimately not impactful.

A logic chain with a strong and convincing rationale and set of objectives has a good chance of being a productive intervention. By contrast one without these elements may run the risk of simply being a busy and noisy one, which delivers a lot of activity but

doesn't really address the core issues in play, because it never understood what these were.

Q: The Interreg programmes are approved. So for the programmes this logic chain is there, either good or bad. So what would be your advice for those programmes which have created a more complicated logic chain or a weak logic chain? What can be done now to maybe make it easier in view of the evaluation?

A (Simon Pringle, evaluation expert): Logic chains and associated theories of change should not be seen as fossils, forever locked at a point in time. Yes, they are developed at a specific point — ideally as part of the programme's design and approval — but they can and should be updated constantly to reflect the changed context for the intervention, the operating experience of the programme in practice, and the lessons that are emerging (e.g. are we actually supporting the right sort of beneficiaries, are we delivering the right mix of projects?).

Practically, therefore, you and those stakeholders with whom you are working need to see the logic chain and TOC as dynamic statements of what you are trying to achieve. So, use them as 'living documents' to reflect accurately reflect what your programme is seeking to achieve, both in their original design and in the light of operating experience. And do past versions of your logic chains and theories of change in a safe place, so that evaluators can see evidentially how you have moved on the content of these as your programme has matured over time.

Q: Interreg programmes are sometimes very small compared to other funds, compared to other policy developments. How can the programme distinguish what are the results of the programme? How can the results be attributed to problem solving, to the impact of the area?

A (Simon Pringle, evaluation expert): Oh that the world was as simple as single funders and single actors for our programmes! But, of course, it isn't, and indeed much of the imperative for the delivering of our programmes in partnership is to secure other inputs and resources, although this of course adds to complexity.

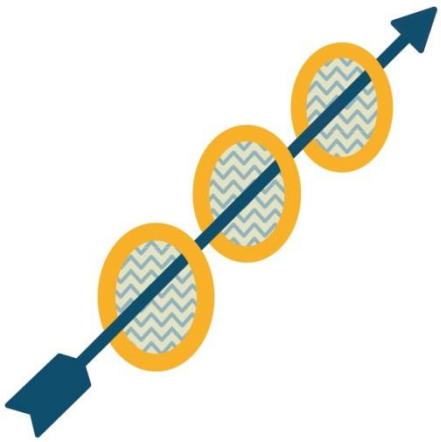
How do we deal with this complexity in terms of the logic chains and theories of change for our programmes? Ideally, whilst we as Programme Managers are responsible for the logic chain and theory of change for our programme, they do need to be understood and owned equally by those partners (and funders) who are working with them, so that we all have a common understanding of what we're trying to do. This common understanding should help deliver alignment in our thinking, activities, and resourcing.

So far, so good, but to the core of your question —how do we identify our share of the impact in a multi-partner driven programme? There are no easy answers for this. My own approach, having been involved in the evaluation arena for 20+ years is to attribute the shares of results/impacts with the shares of input funding made. So, if we have funded a programme 100%, it is reasonable for us to claim 100% of the results and impacts generated. By contrast, if our inputs have been 25% of the total resource made to the programme, then we should claim, on an attributed basis, 25% of the results/impacts for ourselves, with the remaining 75% going to others.

In the past, the Commission has taken the line that it claims 100% of results/impacts from programmes, despite its share of the inputs. Okay, at one level, but it does very quickly give rise to the issue of 'double counting', where the summed claims of programme results/impacts made by partners are very often in excess of what was achieved practically on the ground. Neither helpful nor sensible economically!

Q: Why should the indicators be linked to conditions, considering that we have output and result indicators? Why is the condition so important?

A (Simon Pringle, evaluation expert): Indicators of condition tell us where we are starting from. If they are selected properly, they can be tracked and monitored to show how our programme — which is intending to address these conditions — is performing over time.



Output and results indicators are often helpful in understanding how 'busy' programmes are but without understanding how conditions are (hopefully) improving as a result of our invention. Effective monitoring and evaluation need to check both changes in condition indicators and 'our response' indicators — typically outputs and results.

Q: Should you have a logic chain per specific objective or can you do one for your programme on the whole?

A (Simon Pringle, evaluation expert): You can apply logic chains at all levels. If your programmes' objectives are clearly very different from one of the, then it may be sensible to have a logic chain for each of those thematic areas or objective areas. If it's a more 'homogenous' programme, then it's probably sensible to have a logic chain at the level of the programme overall.

Ultimately, the decision is yours —what you are fundamentally trying to do in creating a logic chain (and subsequent) theory of change is to have a clear and communicable device that allows you and partners to understand, manage, and ultimately account for the performance of your programme. If that means one logic chain or theory of change overall, then fine. Alternatively, if you need three logic chains and three theories of change to capture, say, three different strands within your programme, then equally fine. But avoid having so many logic chains or theory of change depictions that it becomes 'difficult to see the wood for the trees.'

Tip: Building a 'Logic Chain'

- There are no 'right answers' as to what a logic chain should look like – most logic chains should be built around the eight components depicted earlier in this pack, with the associated descriptive narrative simple and clear.

- It does take time to get into the thinking of 'logic chain' and 'theory of change' approaches. But Rome wasn't built in a day, and as you progress and become more exposed to logic chain and theory of change thinking, you will steadily become more expert.

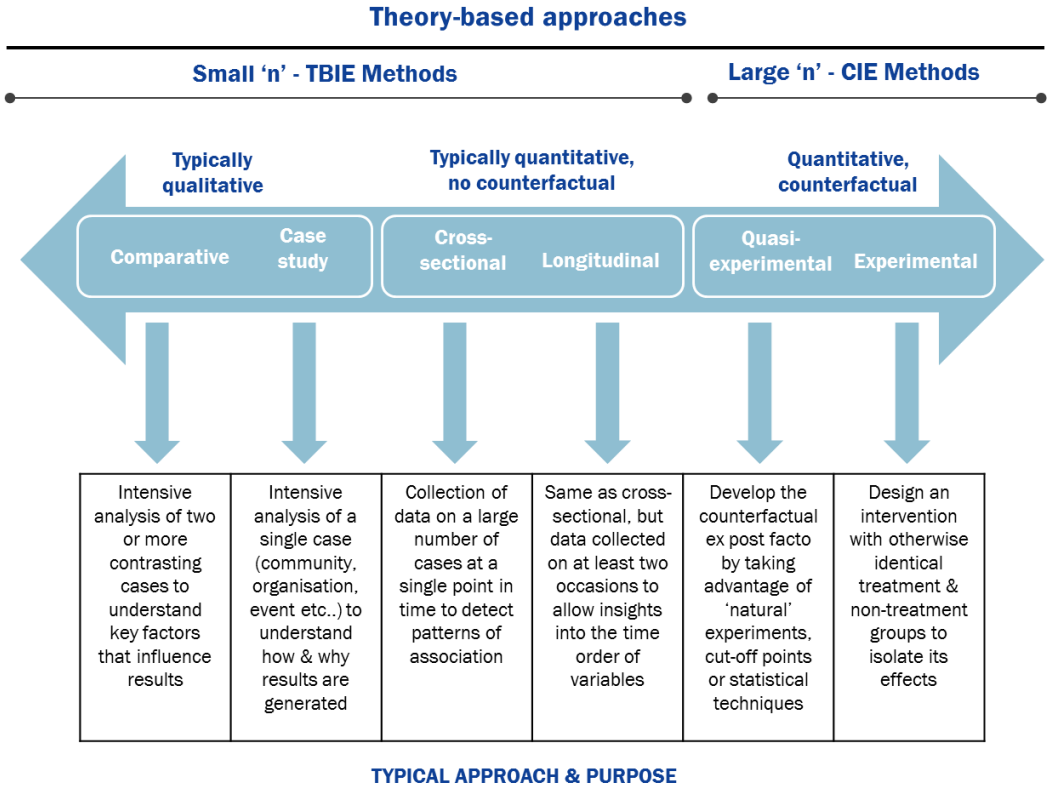
- Do not believe that because your programme has been signed off and there was a logic chain inside it, that it's the world's best logic chain. Every logic chain can be improved and can be developed and can be enhanced in the light of operating experience.

- Not all of you will necessarily have currently a one-page depiction of your logic chain. But as part of the process of securing Commission approval, your programme documents must contain the elements of logic chain thinking. If you don't have a one-page depiction, go back to the office, pull your documents out, and try and produce a simple logic chain aligned with the thinking from today's presentation. If you already have a one-page depiction, then go back to it and see if you can sharpen and deepen it further based on what you have learnt. Remember, a good and strong logic chain is a key input to any evaluation activity.

2. Approaches and Methods of Theory Based Impact Evaluation[3]

Much evaluation practice is now focused on 'theory based' approaches. By theory, we mean those assumptions, hypotheses, and predictions that have been assembled to explain how the intended programme will work.

Drawing on developments in academia (in particular social science and economics) a range of 'theory based' methods now exist for evaluators to deploy and take forward. On the one hand, theory based methods take in full-blown experimental approaches (e.g. Counterfactual Impact Evaluation) where datasets are large (i.e. n=large) and beneficiary populations are homogeneous, through to comparative methods largely based on qualitative approaches. The graphic below gives a sense of the spread of these methods, all of which are based on theory based thinking.



2.1. The Importance of 'n' in Determining Evaluation Methods

A key consideration in choosing the appropriate evaluation method, as is evident from the figure above, is 'n' - that is how large is the population of programme beneficiaries. Ideally, 'n' will be numerically large so that statistically robust and significant methods evaluation methods can be used. Similarly, the population of beneficiaries will, to the maximum extent, be homogeneous, so that the intervention being evaluated is the only

[3] Based on Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012

difference between beneficiaries and their wider populations, with any changes thereby being fully attributable to the programme's intervention.

In the real world, Interreg programme managers face two challenges:

- 'n', the number of beneficiaries (both absolutely and as a percentage of the population from which they are drawn) is often relatively modest. As a consequence of small 'n', statistically resilient and rigorous methods are difficult to progress.
- Homogeneity in beneficiaries is intrinsically difficult for Interreg programmes to achieve, essentially because they operate across different territories, each of which has their own different rules, regulations, cultures etc.

As a consequence, for most Interreg programmes, impact evaluation methods will resolve around so-called small 'n' methods. Like large 'n' methods (such as Counterfactual and full-blown experimental methods) these are based on theory-based approaches, but because they cannot secure statistical definitiveness, they give greater emphasis to qualitative considerations.

These small 'n' and so-called theory-based impact evaluation (TBIE) methods come in a range of flavours – from straightforward reviews of if/how the programme's theory of change has been achieved in practice, through to Realistic Evaluation (where the emphasis is on identifying what combinations of context and mechanism most make for impact), and on to techniques such as the Success Case Method (where the focus is on the naturalistic enquiry of the very best and very worst results of intervention, and the role of contextual factors in driving these).

This variety of TBIE methods is summarised in the slides[4] below. Accompanying this note is a formal paper from White and Philips[5] (part of the International Initiative for Impact Evaluation) which sets out the underlying details of the methods, together with their strengths and benefits. Be aware that the paper is long, but the reading time that is required will be worthwhile.

2.2. Group I: Theory-based Methods to Determine Causes of Observed Effects & how 'Additional' Observed Outcomes Occurred[6]

The Group I approaches (Theory of Change, Realist Evaluation, General Elimination Methodology, Process Tracing, Contribution Analysis) have the goal to explain WHAT has occurred and HOW it has occurred. All four approaches aim **to get an understanding of the causal chain connecting observed outcomes to an intervention**. *"They seek out evidence to substantiate whether a programmes' specified theory of change occurred in practice…"*[7]

[4] This slides were presented at the event Impact Evaluation: Methods and ToR, Amsterdam, 22-23 June 2016. http://www.interact-eu.net/#o=events/impact-evaluations-methods-and-tor Presentation{Theory Based Impact Evaluation Methods/Simon Pringle
[5] http://www.3ieimpact.org/media/filer_public/2012/06/29/working_paper_15.pdf
[6] Detailed explanation on the group 1 approaches can be found in: Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012,p.9-16
[7] Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012,p.15

2.2.1. Theory of Change

- Takes the logic chain for the intervention and develops this in to a predictive and explanatory depiction of what should happen through the intervention.
- Evaluation explores each step of the ToC to understand whether theoretically predicted changes occurred as expected, &/or as result of other external factors.

So the logic model will be developed into a theory of change (see images below).

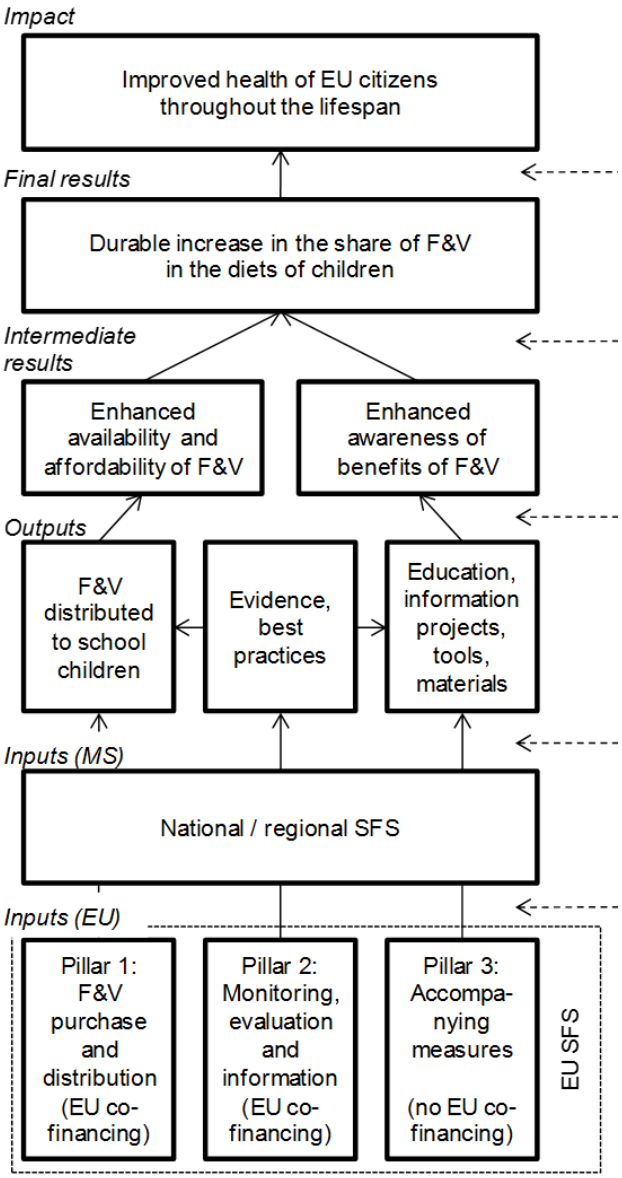


Image: logic model

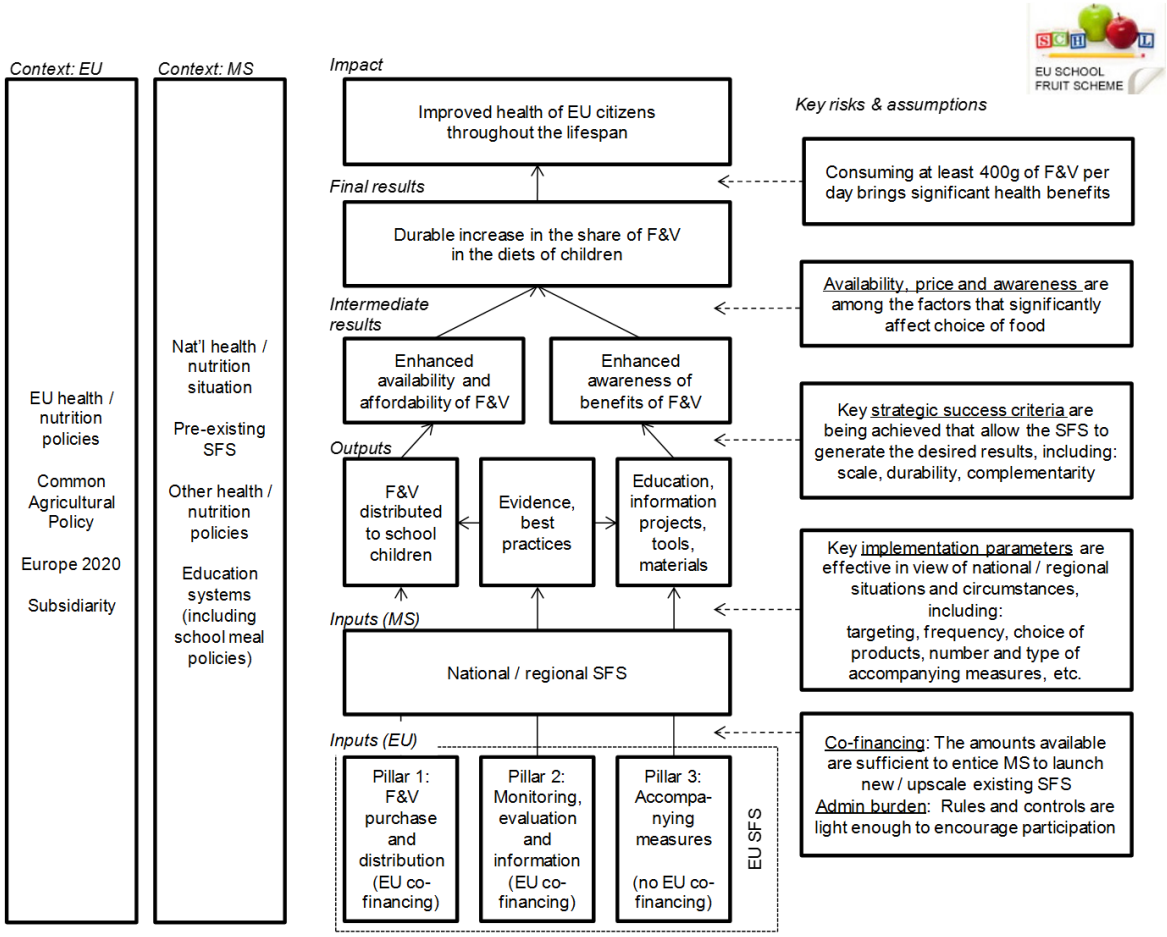


Image: Theory of change

2.2.2. Realist Evaluation

- After Pawson & Tilley (1997)
  - Sceptical about (quasi) experimental approaches
  - "Where several evaluations of similar interventions in different contexts exist, the most usual finding is that the results vary."
- Key concept: context + mechanisms = outcomes
  - What Context Mechanism Outcome Configurations [8](CMOCs) appear the most successful
  - By doing so, RE seeks to understand "what works, how, in which conditions & for whom"
- So, in practice?
  - Realist Evaluation is not a method but a way of thinking, so realist design can be incorporated within almost any evaluation

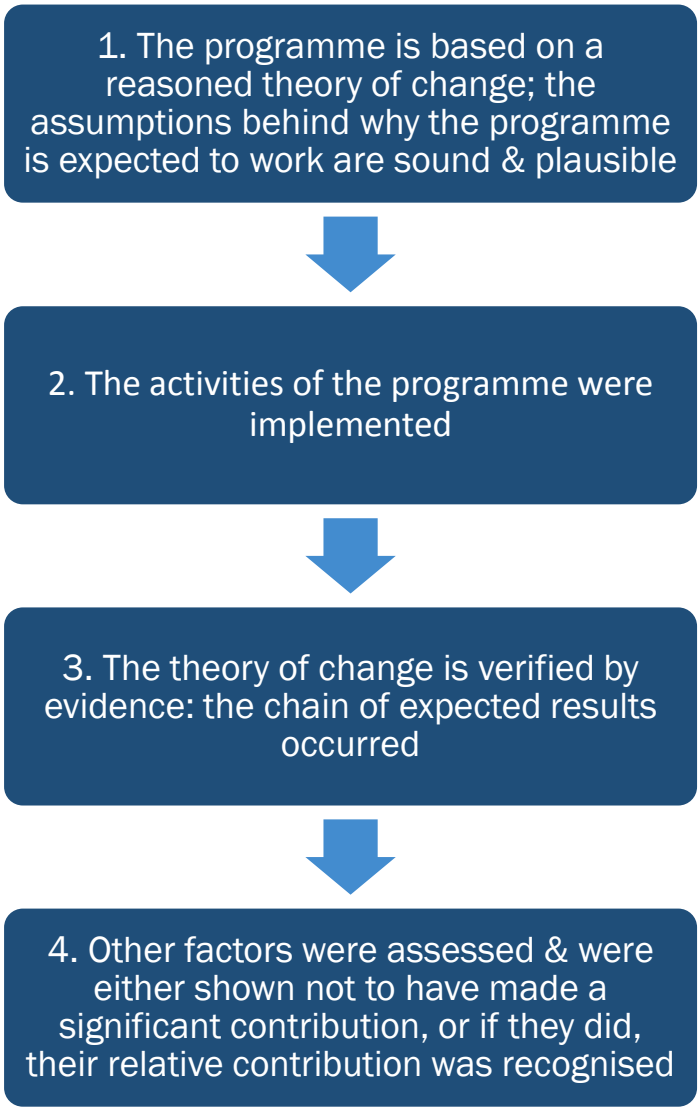
[8] A CMO configuration aims to identify 'What works, for whom, in what respects, to what extent, in what contexts, and how?". In order to answer that question, realist evaluators aim to identify the underlying generative mechanisms that explain 'how' the outcomes were caused and the influence of context. .

Typical steps:

1. Theory & hypothesis formulation
    - Carry out research to establish the prevailing Middle Range Theory[9] (MRT)
    - Map out a series of conjectural mini-theories or CMOCs
2. Data collection
    - Includes quantitative & qualitative research
    - Aim is to refine, refute or demonstrate how CMOCs have operated in practice
3. Data analysis & conclusions
    - How mechanisms have operated in programme contexts to generate results – which CMO configurations were substantiated, which were invalidated, & which need to be revised

2.2.3. Contribution Analysis

- Developed Mayne (1999) to address the problem of attribution.
- Are observed results due to programme activities rather than other factors?
- Conceived as an alternative to experimental designs, when these are not feasible.
- CA sets out to verify the theory of change, but also takes into consideration other factors
- Causality/contribution is inferred 'beyond reasonable doubt' by assessing factors.



[9] Middle-range theory starts with an empirical phenomenon (as opposed to a broad abstract entity like the social system) and abstracts from it to create general statements that can be verified by data

2.3. Group II: Factors Perceived to have been Important in Producing Change, with a Strong Emphasis on Stakeholder Views[10]

In comparison to the Group I approaches the **Group II approaches** (Most Significant Change, Success Case Method, Outcome Mapping, Method for Impact Assessment of Programs and Projects) do not set out to address the attribution of cause and effect but **place the stakeholder participation at the heart of data collection and analysis**. *"They target programme beneficiaries, implementers and other key stakeholders in order to establish what factors are perceived to have been important in producing change; in doing so they aim to gain an insight into the how a programme is performing and the part that it is playing in driving change."*[11]

2.3.1. Most Significant Change (MSC)

- After Davies & Dart (2005).
- Participatory process involving sequential collection of stories of significant change which have occurred as a result of intervention.
- Linked process of sifting by stakeholders to select, discuss, & crystallise most significant changes.
- Typically, "looking back over the last XX, what do you think the MSC in XX or YY has been."
- If done well, can generate useful information for the specification & subsequent assessment of a Theory of Change.

2.3.2. Success Case Methods

- After Brinkerhoff (2003).
- Narrative technique using naturalistic enquiry & case study analysis
- Intended to be quick/simple.
- Focus deliberately on very best & very worst results of intervention, & role of contextual factors in driving this.
- "Searches out & surfaces successes, bringing them to light in persuasive & compelling stories so that they can be weighed… provided as motivating & concrete examples to others, & learned from so that we have a better understanding of why things worked & why they did not."

[10] Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012,p.16-21
[11] Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012,p.16

2.4. Qualitative Comparative Analysis

- Case-based method which identifies different combinations of factors that are critical to a given result, in a given context.
- Not yet widely used in evaluation, provides an innovative way of testing programme theories of change.
- Qualitative data/evidence on potentially relevant causal factors is turned into a quantitative score that can be compared across cases
    - Crisp set QCA: cases coded "0" or "1"
    - Multi-value QCA: allows for some intermediate values (e.g. 0.33 or 0.5)
    - Fuzzy set QCA: allows for coding on a continuous scale anywhere between 0 & 1

Tips: Choosing the Method:

- Let's be honest
    - Evaluation in an EU-funding context still maturing
    - Many of you personally coming new to this area & many of you want to 'do' rather, than 'review'!
    - There's a ferocious market of external 'evaluators'
    - Very significant risk of 'running' before 'walking': that why it is recommended for Interreg programmes – which are not that experienced with impact evaluation approaches- to start with theory of change method (2.2.1.)
    -
- So, four guiding principles to go forward with
    - Be pragmatic – 85% of something is better than 100% of nothing & avoid being overly academic!
    - Be developmental – Rome wasn't built in a day, & evaluative capability needs to build & evolve
    - Become intelligent consumers – think what you are doing, or buying, evaluation-wise
    - Do commit to making time for evaluating things, to build your knowledge

- More information on different evaluation approaches can also be found under: http://betterevaluation.org/approach/

Conclusions TBIE Methods:

- Theory-based impact evaluation cannot rival the rigour with which well-designed counterfactual impact evaluation addresses issues of attribution
- However, done 'right', TBIE can tackle attribution & provide evidence to back up causal claims
- White & Phillips have identified the following *"common steps for causal inference in small 'n' cases":*
    1. Set out the attribution question(s)
    2. Set out the programme's theory of change
    3. Develop an evaluation plan for data collection & analysis
    4. Identify alternative causal hypotheses
    5. Use evidence to verify the causal chain

2.5. Evaluation Techniques[12]

Evaluation Techniques are techniques which could be used in all the approaches above.

2.5.1. Contextual & Documentary Review

- *What did we think we were doing?'*
- Desk-based review of
    - The problem/challenge faced (context – data)
    - The case for intervention (rationale – arguments)
    - Our practical commitment (objectives, inputs)
    - Progress so far (activities, outputs, & processes)
    - Knowing what we know now:
        - How logically consistent is all of this?
        - Do we need to change track?
- Sources: secondary data (local, national, European) original programme documents, application forms, appraisals, approvals, monitoring data & reports etc.

[12] The explanation of the evaluation techniques is based on the slides, which were presented at the event Impact Evaluation: Methods and ToR, Amsterdam, 22-23 June 2016. http://www.interact-eu.net/#o=events/impact-evaluations-methods-and-tor Presentation{Theory Based Impact Evaluation Methods/Simon Pringle

2.5.2. One-to-one Consultation

- [As an informed viewer] *'What are you observing about the intervention?'*
- Detailed consultations with key stakeholders
    - Policymakers/funders
    - Adjacent programmes
    - Delivery bodies
- Modes: face-to-face – telecom – postal – online
- Useful for scoping the issues & for cross-checking messages from elsewhere in study

2.5.3. Surveys

- [As someone who is impacted] *'What has your experience of this intervention been'*
- 2 groups
    - Beneficiaries – intended or otherwise
    - Non-Beneficiaries – typically those who were ruled out
- Modes: Face-to-face – telecom – postal – online
- Typically, self-reported view & observation
- Prone to
    - Last event bias
    - Memory decay
- Questionnaire design & analyse-ability a key challenge

2.5.4. One-to-Many Consultation

- [As informed viewers] *'What are you observing about the intervention?'*
- Similar to one-to-one consultations, but multi- rather than bi-lateral
- Efficient to setup/deliver
- Prone to
    - Superficiality
    - Herd effects
    - Loudest voices
    - and tend to be primarily qualitative in observation
- Often useful to calibrate headlines from 1to1 consultations & surveys

2.5.5. Case Studies

- *What has worked well & less well*
- Provide deep-dives into specific aspects of the intervention
    - Process
    - Impact
    - Learning
- Typically, done face-to-face – so, resource intensive
- Judgement required to establish rounded view
- Can be hard to secure consensus amongst consultees
- Can be difficult to synthesise findings across case study authors

2.5.6. Learning Diaries

- *Real-time recording of intervention experiences*
- *Avoids memory decay & last-event bias*
- *Does require discipline on part of participants to maintain diary*
- *Need recording interval that makes sense – related to speed of changes happening/progress being achieved*
- *Helpful to frame wider consultation/survey work*

Q: Does theory of change offer us information on everything that we would like to evaluate in our programme? Does the theory of change show us the extra impact?

A (Simon Pringle, evaluation expert): The theory of change provides the assumptions and hypotheses that a formal evaluation would test and explore. Impact will be one such area, but there are often others. If the theory of change is not accurate, definitive, or up-to-date, then there is a good likelihood that the evaluation's findings will be similar. This reinforces the point made earlier about the need for logic chains and theories of change to be 'living documents', which remain accurate and up-to-date at all points. My own experience in undertaking of evaluations is that the more robust and accurate the content of the theory of change is, the better, stronger and administratively easier the process of evaluation — by yourselves and/or external contractors — will prove to be.

Q: Is the theory of change also rooted in all the other methods?

A (Simon Pringle, evaluation expert): Most current evaluation methods are 'theory based', although within this the methods vary from being very numbers-driven e.g. Counterfactual Impact Evaluation through to the more comparative qualitative methods.
The precise method you adopt depends on your need and context —e.g. CIE is the most statistically robust method available to us, but large beneficiary populations are needed, and these need to be homogeneous. Both are very difficult requirements to achieve within the context of Interreg programmes. By contrast, the qualitative methods don't require large populations and/or homogeneity, but they are not very informative in terms of quantitative impact.

The key issue the Programme Managers have in progressing their impact evaluations is to pick the 'right' method for the perspective you want to understand – in a very real sense choosing your (theory based) method depends on the context and evaluation imperatives which we are facing.

Tips: Link Ex-ante Evaluation and Theory of Change

Go back and look at your ex-ante evaluation in particular, see if you can find the bit of the ex-ante evaluation that explicitly talks about a theory of change. If there is, think about developing, think about moving it further, and if you have an ex-ante evaluation that doesn't explicitly talk about or present a theory of change, then take the opportunity, perhaps using a template (e.g. as above), to develop one. So, you begin to tie it together, the ex-ante intelligence with the theory of change.

Q: What is the link of theory of change and evaluation?

A (Simon Pringle, evaluation expert): A good evaluation will be focused on your theory of change, and will through a process of deploying different tasks and approaches e.g. surveys, partner interviews, document review, focus groups et cetera) explore, test, and assess whether the assumptions and predictions in your theory of change were valid, and remain so currently.

All evaluators will seek a current and robust theory of change as the starting point for evaluation activity.

Q: If the evaluation question isn't answered and you think there are disparities do we go back to those theories of changes and sort of change them to the way that we think the evaluations, it needs to be aligned?

A (Simon Pringle, evaluation expert): As commented elsewhere, the logic chain and theory of change should be treated as 'living documents', which are revised and updated in the light of experience and learning.

Sometimes the assumptions we made ex-ante are not realised, and we need to update the theory of change to reflect this. But don't forget in so doing that the theory of change is not a justification for the programme —rather, it should be the other way around i.e., the theory of change tries to explain and predict how, why, and when the programme will work to address the arguments set out in the condition data, the rationale for intervention, and the objectives.

Tips: How to get a better Impact Evaluation

- Do commit to thinking hard about developing a theory of change both for your own understanding but also to prepare the way for your impact evaluations. Start with developing a clear and simple logic chain, and then develop this into a descriptive and predictive theory of change. Once you are happy with your theory of change, then begin to think how in evaluation terms the various assumptions and hypotheses which the theory of change contains would best be tested, proved and measured through evaluation activity. This might involve using techniques such as realist evaluation, contribution analysis etc.

- Do think very carefully when you see proposals from external evaluators that they are giving you what you think are appropriate, mixed, cost-efficient methods for the evaluations you want to progress. Have they understood the logic chain of your programme, do they appreciate the assumptions and hypotheses that your programme's theory of change has set out, and are the tools, techniques, and approaches that they are advocating sensibly aligned with your evaluation requirements and how your programme has operated?

- Don't hesitate to challenge the external evaluators both at proposal stage and after selection – 'why are you saying that we do it that way, where is the learning for this programme from other evaluation activities, is that method feasible given where our program currently is at' etc. Focus on being an intelligent consumer of external evaluation input

2.6. Evaluation Questions

Q: In order to do an evaluation the programmes need some evaluation questions on what you want to evaluate. The more the evaluation questions are focused, the more the Interreg programmes can assess the real impact. What do you think?

A (Simon Pringle, evaluation expert): As previously stated, the clearer and more robust the logic chain and theory of change for your programme, the easier impact evaluation should (although not always!) be.

On questions to ask, these are typically framed by what your evaluation is trying to do. So, a process evaluation will focus on probing processes, procedures and behaviours, whilst a quantitative impact evaluation will want to dig for harder numerical results and outcomes.

What the theory of change enables you to do is to identify the key assumptions that the programme is working to, and these can then be used to frame the questions that the evaluation's purpose is seeking to address. So, for a process evaluation, 'are the

beneficiaries we have approached the right ones, did the partnerships intended added value deliver etc; similarly for a quantitative impact evaluation 'how have the conditions changed, has our objective in moving from X to Y been achieved etc.'

Remember, each evaluation is its own exercise, so I'd want to discourage you from having overly standard lists of questions that you can simply pick and mix from when evaluators are selected. You need to be more reflective than simply relying on standard questions drafted elsewhere!

Q: How much should programmes prescribe what evaluation questions are versus letting the evaluator come up with it? If the programme has a pool of questions and now they need to narrow it down and prioritise and to understand what is realistic to get an answer from? Does the programme give some examples or does the programme let the evaluator come up with their input?

A (Simon Pringle, evaluation expert): There are two points here:

- If you are using an external evaluator, then one of the reasons for so doing is to access their expertise and knowledge. You need to work and exploit this so that your evaluators are thinking hard and creatively for you. Ideally, they will have good past experience to draw on, and you should work to exploit this as part of their evaluation work. But remember that external evaluators are your agent – they are working with and for you; your role is to get the best out of them, not simply abdicate all responsibility for the evaluation to them so that you can tick the 'evaluation done' box on your management schedule. That sort of evaluation activity delivers you no learning or internalisation of evaluation knowledge, and as such should be avoided.
- Second, you will know your programme better than anyone. You may well have designed it, implemented it, and will probably have experience of operating it. You will, therefore, have considerable learning and experience to bring to the evaluation exercise. So make sure you are open to sharing your experience, offer your full and frank opinions as to what has gone well or otherwise, and bring your knowledge alongside that of the evaluator's technical expertise. It's this combination of practical understanding and technical expertise that gives very strong and powerful evaluation results.

2.7. Learning from Evaluations

Q: Are there any good practices or examples of how we can learn from evaluations e.g. the PMC can make decisions based on them or the JS, the beneficiaries? It is more than a document; it is something which is getting incorporated into practice.

A (Simon Pringle, evaluation expert): A couple of points here:

- Reading good evaluation reports from elsewhere is often a useful technique for building your own knowledge and understanding of evaluation tools and techniques. Interact may well have a role to play here in providing a reading library of such exemplary evaluations.
- All evaluations have different layers of audience – programme staff, PMCs, wider partners etc. Good evaluation recognises these different audiences, and is undertaken and reports in ways that recognises the different needs and wants of the different groups. A key distinction in good evaluation is distinguishing between 'strategic' and 'tactical' findings and recommendations, and making sure these are played out to the appropriate audiences.
- There is a temptation when using external evaluators to abdicate responsibility to them. In my experience, the best evaluations are undertaken as a partnership exercise, using the combined expertise and knowledge of programme staff with the technical abilities are skilled evaluators. This partnership approach is the best way of delivering synergy and success.
- A final point, drawing on my own style of evaluation. 'Surprised' clients are often unpredictable and unhappy ones. Therefore, a really important technique in progressing evaluations are to make sure that the messages and themes, for good or ill, are played out early and promptly, so people are aware in general terms of what is coming, and don't end up being surprised at the end of the exercise.

Q: In our evaluation plan we have two areas which we will look at: one is the operational side and then we have the programme priorities. Our argument is to separate operational and programme specific ones. How can we make sure not to have a fragmentation in our evaluation?

A (Simon Pringle, evaluation expert):
Programme fragmentation is always tricky, and particularly so when you are engaging in an evaluation activity.

In a very real sense, how you decide to assess aspects of your programme for evaluation purposes is down to you, but as a broad rule I would counsel against having too many different logic chains or theories of change for your programme. Ideally what you want to have is a single integrated logic chain and a single integrated theory of change which spans the spectrum of operational and strategic issues, and you then use the process of evaluation to drill down into aspects of operational and aspects of strategic activity as appropriate. In this sense evaluation activity becomes a bit like examining the two sides of the same coin —different, but intimately linked by a shared and common logic chain and theory of change.

Q: We are asking ourselves if we should go for a bottom up approach or for a top down approach when we are designing our interventions. We have many regions in our programme. Even if we have a very good knowledge of the programme area we can never have the same as the people working on the ground in the regions. That is why we need to leave a certain level of flexibility and freedom to the beneficiaries to come up with their own ideas. And then when we receive various project applications: some choosing topics that may seem not so attractive. In this sense we have a very limited capacity to steer in which directions the project will be going. This problematic also reflects then the impact and the impact evaluation which will be done at the end. So some topics might be very well covered and others might be tackled in a marginal way. And this then is not so much related to the success of a programme but reflects more the real needs which are there on the ground.

A (Simon Pringle, evaluation expert):
Part of the way we build in innovation to our programmes is by bringing together our own (programme-led) top-down thinking with the bottom-up knowledge and expertise of our national/local partners. This coming together can be very rewarding and fruitful, and vital in ensuring that capacity builds both within programmes and amongst our partners.

This is where a good and clear logic chain and theory of change can pay dividends. The theory of change in particular sets out programme assumptions, hypotheses, and expectations, around which programme managers can design and configure Programme Calls, top down. But a good theory of change suitably shared and promoted can also be used by national/local partners to help bring forward innovations and new approaches within the scope of the programme's objectives. In this sense, a good clear and robust theory of change can be an important device for linking together top-down and bottom-up expertise, and delivering real synergy.

Tips: Learning from Evaluations

- If you made a lot of evaluations over the year and you want to be able to conclude and learn from it, it would make sense to have a synthesis at the end of the period which looks at all your evaluations and your conclusions, because the different evaluations look at different aspects.

- The programme does not have to do separate evaluations on different Specific Objectives. The programme can make just one evaluation covering many Specific Objectives.

- There will be some Specific Objectives where you won't be able to do an impact evaluation, because of the limited number of projects: there maybe you just look at the result indicator and the value before and after the intervention. Based on the number of projects and based on the implementation you might be able to say more. Then you can draw conclusions from it. The role of evaluation is to have some policy learning: some of the intervention might not be designed well in order to achieve the Objective - maybe it will be a different objective which you will be achieving with this intervention. What the EC is pushing on is a policy learning effect and it is not the fact that the programmes have an impact evaluation on each single Specific Objective.

- Impact evaluation is a very wide term. The Evaluation Unit defined impact evaluation in the guidance document and considers a counterfactual or a theory-based evaluation as a proper impact evaluation. The regulation says you need to look at the impact. But the impact might be simply just the change of one single indicator for which you don't have to do a proper evaluation. You just have to look at the indicators.

- Without understanding your intervention you won't be able to do a proper evaluation. If you have a bad intervention you can spend as much money, but there will be no results: if you already know in advance that your intervention is not conclusive and not reaching your objectives there is no point of setting up your ToR and implementing your evaluation. Then you have to focus in a different way, e.g. maybe you could then focus only on a few projects. That is why your evaluation plans are not carved in stone. The EC foresees that programmes review evaluation plans on an annual basis and if necessary adopt them.

2.8. Added Value of Cooperation

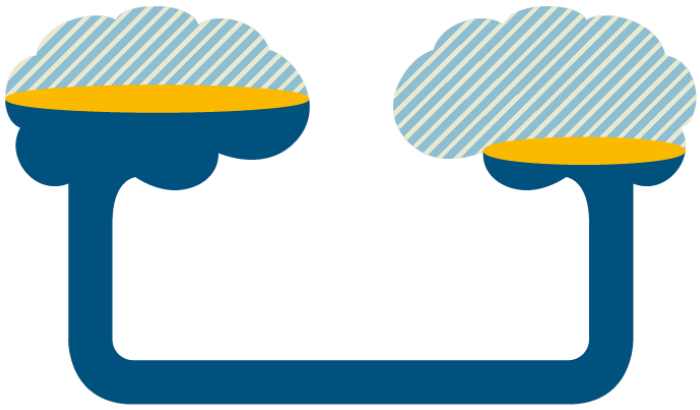
A key aspect of Interreg's work is to bring together partners across different jurisdictions to build common understanding and approaches.

This cooperation activity is often overlooked in evaluation work, and going forward more attention should be given to it, especially by Interreg programmes as they look to better report on their results.

In the UK, work around cooperation has identified typically five types of so called strategic added value, which are as follows:

- Strategic leadership and catalyst: Articulating and communicating development needs in the programme area, opportunities and solutions to partners and solutions to partners and stakeholders in the programme area and elsewhere
- Strategic influence: Carrying-out or stimulating activity that defines the distinctive roles of partners, gets them to commit to shared strategic objectives and to behave and allocate their resources accordingly
- Leverage: Providing/securing financial and other incentives to mobilise partner and stakeholder resources – equipment and people, as well as funding
- Synergy: Using organisational capacity, knowledge and expertise to improve information exchange and knowledge transfer and coordination and/or integration of the design and delivery of interventions between partners
- Finally, Engagement: Setting-up the mechanisms and incentives for the more effective and deliberative engagement of stakeholders in the design and delivery of programme emphases.

As the Interreg programmes look to plan and progress their impact evaluations, paying attention to these aspects of strategic added value should be a core element of their thinking.



3. Counterfactual Impact evaluation (CIE)

CIE is often argued to be the 'gold standard' in evaluation. It is statistically robust and resilient and if done well, with appropriate control groups, is capable of isolating the particular effects of the programme intervention.

For all its quantitative strength however, CIE is weak at explaining the story of programme performance and process. As such, CIE is often undertaken alongside other more quantitative methods, with the results of quantitative and qualitative approaches being synthesised together to form an integrated evaluation assessment.

What are the requirements for CIE? There are essentially two:

- first, a large number of beneficiaries to allow statistically rigorous and significant analysis is needed. In many cases, the population sizes required by CIE will be many times greater than the populations that INTERREG programmes are working with routinely, and from this perspective alone CIE may therefore not be a primary method for Interreg impact evaluation.
- second, CIE requires a high degree of homogeneity in the composition of beneficiaries and their characteristics, so that the 'treatment effect' of the programme intervention can be easily isolated. This again is somewhat problematic for Interreg programmes, given that by their very nature they are transnational, and as such immediate heterogeneity is introduced.

This recognised the following points applied to CIE methods:

- CIE relies on statistical tests of significance between treatment & comparison groups. This requires a large 'n' (sample size)
- CIE is a serious option where: the total population is large; the treatment group; the treatment itself, and/or the wider context, is homogeneous; the intervention affects very defined segments; budgetary, political or other constraints allow a sufficient sample size or use of comparison group.

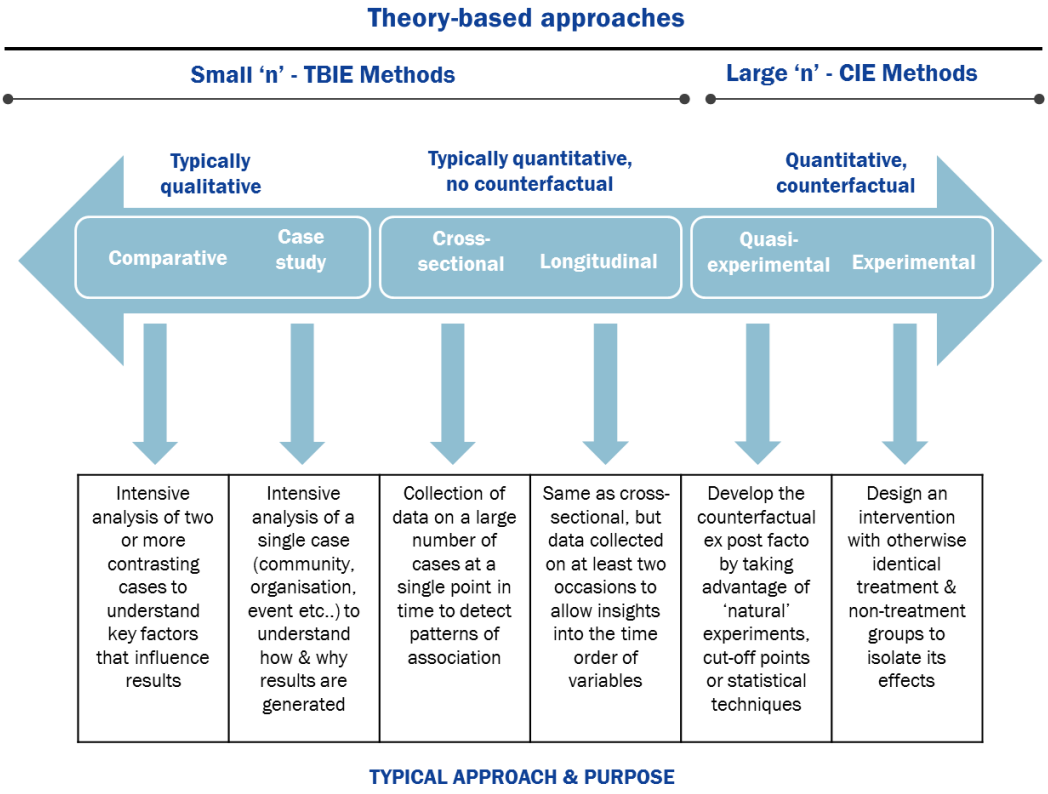
Conclusions

- Theory-based thinking still holds but CIE activity used to deliver the statistically robust impact assessment
- Methods used similar to TBIE
  - Doc review, consultations, case studies, learning diaries etc.
- But quantitative impact story delivered by formal CIE analysis
- CIE, & its underpinning statistics, are specialist areas
  - Largely delivered by academics
  - Not always easy to work with
  - Demanding to follow what is going on – 'Black Box'

4. Counterfactual Impact Evaluation (CIE) versus Theory Based Impact Evaluation (TBIE)

Which, then, is the preferred way forward – CIE or TBIE?

In practice, the dichotomy is a false one; both approaches have their strengths, but each has their own requirements, limitations, and expertise demands. Ideally, a mixed methods approach would combine the benefits of each. However, given the demands of CIE for large and homogenous populations, coupled with a building level of evaluation expertise generally amongst the INTERREG programme community, it is likely that for the next two or three years, TBIE will be the more feasible way forward for impact evaluation activity. This is not to say that CIE will never occur, but is unlikely for this period to be the dominant methodological approach.



Conclusions based on the discussion of the event[13]	
Counterfactual Impact Evaluation (CIE)	**Theory Based Impact Evaluation (TBIE)**
PRO + Fashionable + Most objective approach if it works properly (homogeneous intervention with enough repetitions)	**PRO** + Cheaper + More pragmatic + Feasible for Interreg programmes + Produce good narratives (a good qualitative perspective)
CON - Expensive - Black Box - Huge number of intervention needed - Huge number of repetitions on a consistent basis needed	**CON** - More subjective than CIE - Robustness of quantitative data is low - Dependent on a solid logic chain and theory of change

[13] Impact Evaluation: Methods and ToR, September 2016, Amsterdam

5. Budget, Data

5.1. Budget

Q: What is the formal budget for your impact evaluations in cooperation programmes?

A (Simon Pringle, evaluation expert): I typically work in environment where one to one and a half percent of programme cost for evaluation is calculated. But avoid adopting an overly mathematical mind-set - both internally-driven or externally-commissioned evaluations cost money, and to be done well, they need to be appropriately resourced. And a small innovative programme which is doing pathfinding work for the first time may indeed have a higher evaluation budget than a larger programme which is delivering services with little innovation and novelty for the nth time. Again, it's a question of what internal knowledge on evaluation you have available…
All this said, evaluation is probably one activity you don't want to be too economic with.

Experiences from the programmes participating in the event:

- Most Interreg programmes (participating in the Amsterdam event) use a very small percentage (a lot less than one percent of their programme budget) for (impact) evaluations.
- It's quite interesting how .wide the range of money spent on evaluation is.

5.2. Data

Q: Our programme has a pool of questions that we want to be answered.- We have indicated these in our evaluation plan. Now we have to sit down together with the evaluation steering group to prioritise which questions we want to focus on. So, can we even define what data we need for before you define the questions you want to answer?

A (Simon Pringle, evaluation expert): it is almost impossible to undertake effective monitoring (and evaluation) without good quality and regularised data. Moreover, disrupted data flows e.g., data not being collected from the start of the programme, or datasets being changed midway through can mean programme data quickly becomes discontinuous. Data discontinuity is the curse of effective monitoring and evaluation. This therefore points to three imperatives:

- Monitoring data, which allows the programme's logic chain and theory of change to be tracked, needs definition at the outset
- Processes need to be put in place to ensure these data are collected – from beneficiaries, partners, or indeed wider statistical sources - as soon as the programme commences
- data provided needs to be reviewed regularly for quality and consistency, to ensure that a robust and verified dataset is building.

And if beneficiaries are reluctant to provide data, well there's a simple trade – the programme gives the project the money, the project gives the programme the appropriate data.

Q: We are hoping to develop a monitoring system to collect the data that we need. But we still need to be able to define that, and I don't know, are we able to do that before we have decided what our priorities are in terms of what we want to get out of the evaluation.

A (Simon Pringle, evaluation expert): There are two separate issues here:

- What data do we need to collect *routinely* as part of monitoring the programme and its theory of change? In practice, this definition of data should be underway for all the Interreg programmes, given that they are now largely approved.
- Second, with *impact evaluation specifically in mind*, programme managers need to decide what other datasets they need to cover and address. In practice, if the definition of routine monitoring data has been done effectively and is aligned with the theory of change, most of what you need to collect data-wise for evaluation will already be in place.

Q: How can the programme be sure to have enough monitoring data which is of high quality?

A (Simon Pringle, evaluation expert): Four points here:

- As a good programme manager, you will have identified that monitoring data by which you will track and monitor the programme against its theory of change.
- Ideally, this data (because they drive out of the theory of change for the programme) will provide you with the data set around which evaluation can and should focus.
- You will have a solid base of evidence on the programme's conditions from the ex-ante appraisal, which sets out where you are starting from. You may as programmes choose to refresh this in real time (e.g. 18 or 24 months) or you might specifically undertake refreshing this baseline as a key task in your impact evaluation.
- If you do at any stage get stuck in defining your datasets, don't forget to also have wider Technical Assistance resources to draw down to fill specific holes and deal with particular challenges.

The key thing to hold onto is that collecting data from the start is always a lot easier than having to collect historical data going backwards. Nobody – in the programme, as a beneficiary, or as an evaluator - enjoys doing that.

Q: Many of us are now starting to receive a first progress report from projects. Are there things we should think of now already for the format of these progress reports to ask projects to have good, usable evidence on the impact of the projects themselves in order to have good data for the impact evaluation later on?

Feedback from the Interreg programmes: I think cross-checking is very important. Maybe what we will also do is if we have some projects that are quite similar to each other, we try to work them into the same methodologies for evaluating their own impact, so at least we have comparable data.

Feedback from the Interreg programmes: In our programme, we make sure that what the projects deliver is in terms of output aligned with our definitions of the output indicators, to ensure that the projects fit in the programme and vice versa. The monitoring is another source you can ask for additional information - however the beneficiaries need to be prepared…

Tips: Data Collection

- Do not forget from the beginning to start collecting data. If you don't collect the data, you are in trouble. It is too late to wait for the evaluators to give us more information

- If you are not collecting the data from the get-go on your interventions, it will become very difficult indeed to fill that gap. And you are almost destined then in evaluation terms to underreport what has actually happened. And that will just put more pressure and make life more difficult for you. It just means it becomes harder to defend our programmes to the body. So, do think hard about what the data requirements of your programmes are going to be.

- Define your data by which you want to be evaluated, and agree with your partners that they supply that data.

- Have a clear idea what are the data the programme need to collect, how frequently, in what form, and what data do the beneficiaries need to deliver, what are the programmes assumptions, etc.

- Make sure you collect good quality data: I have seen some fantastically designed project monitoring systems, absolutely full of rubbish data. I was evaluating a programme recently with European money which consistently on its targets was overachieving by two million percent on its target. Mind you, that is impressive programme or two million percent impressive programme with a team of supermen or superwomen or a very poorly specified programme in the outset because they didn't really understand its target and its competence. So, data matters. Commit to data.

- You ought to know after projects have been completed, what you have achieved: not only in terms of outputs, which is usually not a problem, but in terms of results and impacts. So, we when you design the programmes - small or large- make it a condition; otherwise don't give the taxpayers' money to the beneficiaries.

- ESPON can support Interreg programmes with your data collection too. Interreg programmes can submit a request for help with their data problem to ESPON. ESPON is selecting two times a year a different topic and issues. So Interreg programmes can send your request to ESPON and they might select their topic and support the programmes with their data request.
- ESPON is also planning to do monitoring systems for the macro-regional areas, so there could be a continuous monitoring system of certain core indicators.

6. Terms of References

Tips: Putting out Term of References for the External Evaluator:

- Don't overly prescribe
- Be very clear what you want the evaluator to do
  - what is the scope of the work,
  - when do you want them to do it,
  - be specific on when you want certain reports delivered
  - communicate any ideas on how you like them to do it
- Come across as an intelligent and informed organisation looking for consultants
- Give an indication for budget: don't give a precise number, because then everybody will bid up to that number; but give a scale of intent (e.g. €100 000 - €150 000); so that the consultant will know the range (if you want one survey or ten surveys ; one stakeholder workshop or ten workshops); give a guide price not exact price
- Do encourage in the brief the provider to think about the skills and expertise that might be necessary, because the sort of programmes you are doing are quite wide ranging and often one single provider will be able to cover all that; so think about affiliations and partnerships: be very clear who in the team you are buying and what level of resource and what input you are getting: break down the tasks, by people and cost over time (e.g. if you get 1000 days and 997 junior result consultant and only 3 days of the director - you definitely got the wrong mix
- It takes up quite a lot of internal resources to manage impact evaluations. So, when you do the timing of the evaluations, it's not only a question which interventions you will cover, but you also need to consider which internal resources you have available.
- You don't have to do impact evaluation on every specific objective. However, every specific objective has to be covered by evaluations and the programmes have to decide how they want to cover it. For example, if there are not many projects in a specific objective implemented, there is no point of investing too much time in it. When you do your evaluation plan and you need to plan your evaluation in a way to justify how you spend the money: e.g. you can either do five small evaluations for five subsequent activities or you can do one big evaluation.
- You can decide to do one framework contract for all evaluations or to do a contract for each single evaluation. Each approach has advantages and disadvantages:
  - small contracts are easier to control. But it might need much work and resources if you have to do a tender each time.
  - The advantage of a bigger contract is that one single person takes responsibility of all the steps. However, it might be more difficult to find an evaluator to cover all the different thematic fields.

Q: What can we programme do in order to find potential contractors?

Feedback from the Interreg programmes: What we do is that we set up a dedicated website section on our programme website where whatever calls from tenders or from project partners is published – we even do it volunteer-based for small contracts, and there is quite a response. So, this could also be an option, that you use this website because the community looks where potential contracts could come.

Tips

In order to find potential contractors you could contact the evaluation societies in the different countries. These evaluation societies could point you to different evaluators.

7. Link Operational Evaluations – Impact Evaluations

This chapter will be further elaborated after the Interact event "Link operational evaluation-impact evaluation, planned for June 2017."

Tips: Links between Operational Evaluations and Impact Evaluations

- You could use the Operational Evaluation as a preparation for the impact evaluation: in the operational evaluation you could have a look at the thematic achievements and define where you have a critical mass of achievement to go deeper into an impact evaluation.
- If the operational evaluation shows that you have limited number of projects in one investment priority that is enough evidence to justify why you might not need to do an impact evaluation, e.g. IP transport if you have a very limited number of projects. The question that you could ask then is why there are not many projects. And then you could reflect if you need this intervention or not?

8. Interact provides …

Q: Is there some way in which Interact performs some quality checks? We are all different programmes and we all do evaluations in a completely different way: some go very much into details. Is there something to maintain a certain standard?

Interact can provide:

- A platform (Basecamp Group Result and Evaluation), where we share different evaluation examples and circulate evaluation material which is interesting (e.g. Interreg evaluation plans, Interreg evaluation reports; guidance documents, etc.); the aim is to learn from each other and collectively and individually get professionally better.
  Please do share your evaluation reports and evaluation plans with Interact so that we can analyse it and also share it on the platform.
- Videos of important presentations (e.g. methods of impact evaluation, ToR, self-evaluation of projects, etc.).
- Q&A documents: guidance and lessons learned from Interact evaluation events.
- Interact is also starting a huge capitalisation effort. Interact is going to work in the next couple of years on many different topics and activities. All this information can be found on the website (capitalisation plan, capitalisation strategies). Interact is starting with the topics of transport, then we are going to move on to climate change.

Annex 1 Acronyms

CIE	Counterfactual Impact Evaluation
CP	Cooperation Programme
CPR	Common Provision Regulation
DG REGIO	Directorate-General for Regional and Urban Policy
ERDF	European Regional Development Fund
ETC	European Territorial Cooperation
EC	European Commission
GDP	Gross domestic product
IP	Investment Priorities
MA	Managing Authorities
MC	Monitoring Committee
SFC	System for Fund Management in the European Union
SME	Small and medium-sized enterprises
SWOT	Strengths, Weaknesses, Opportunities and Threats
TA	Technical assistance
TBIE	Theory Based Impact Evaluation
ToC	Theory of Change
ToR	Terms of References

Annex 2
Exercise: *EuroHungHo (part1): Logic Chain and Theory of Change*

- *Purpose*
  - *Develop Logic Chain & Theory of Change for EuroHungHo*
- *Context*
  - *EuroHungHo – a fabricated project!*
  - *'Improving existing & developing new innovation support services, with a focus on the sectors of special interest to the Programme Area'*
  - *8 countries*
  - *5 sectors of special interest*
  - *Identify/developing R&D projects, pilots/prototypes, demonstrators*

Exercise 1 – completing the logic chain

- *Task*
  - *Using template, develop*
    1. *Logic chain (descriptive)*
    2. *Theory of Change (explanatory & predictive)*

- Defining 'Activities' is easy, with 'Outputs' being a bit harder. 'Conditions', 'Rationales', 'Objectives', and 'Results' are always more difficult to define, and take much more thinking time and brain power. But because they are harder does not in any way mean that they should not be as fully addressed.
- Avoid the temptation of choosing too general objectives. Specific, Measurable, Achievable, Realistic, and Timed should be the key features of all objectives that we define for our programmes. If we cannot define our objectives SMARTly, that tends to suggest we are not really clear about what we are to achieve. And, remember, all good evaluation will as its starting point take programme objectives as its first input and work through these to understand whether what you said you were going to do has, in practice, been achieved.
- Although the logic chain is presented as a linear flow, we should in practice think about it as a 'closed loop'. In this sense, when we are thinking about our programme 'Outputs' and 'Results', we need to be focusing on how these will help address the 'Conditions' and 'Rationale' for our programme intervention in the first place. As Simon says, the biggest testament to our programmes' success would be that they are no longer needed because we have put right all the wrongs that they sought to address.

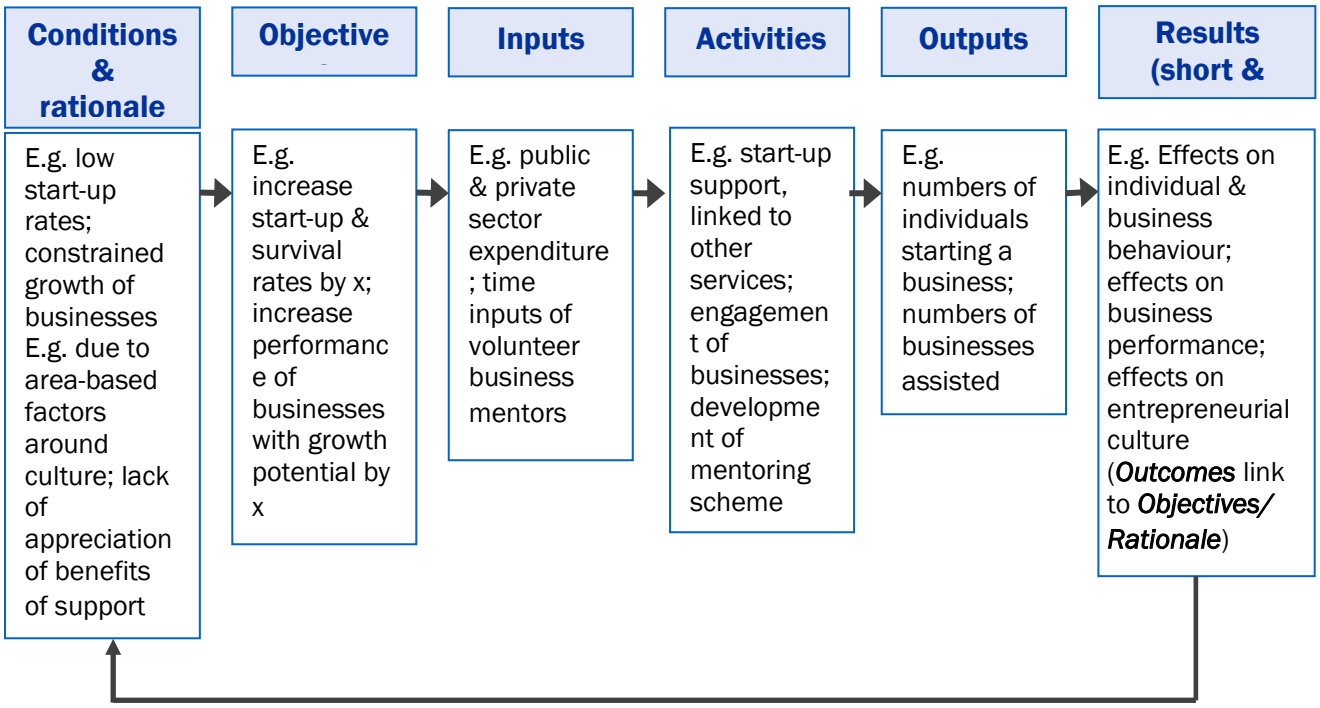
EuroHungHo (part2): Planning for an Impact Evaluation

- *So, what might an Impact Evaluation Plan for EuroGungho look like?*
- *Using presented methods & techniques etc.:*
    - *'What, where & how' of an outline impact evaluation plan*
        - *Which theory based approach?*
        - *What mix of techniques to progress, & sequencing?*
        - *Do as a simple block diagram – template provided*
    - *What pre-requisites*
    - *Timing of impact evaluation activity*
        - *When, & why?*
    - *Resourcing*
        - *What cost to undertake - €s, internal vs external?*

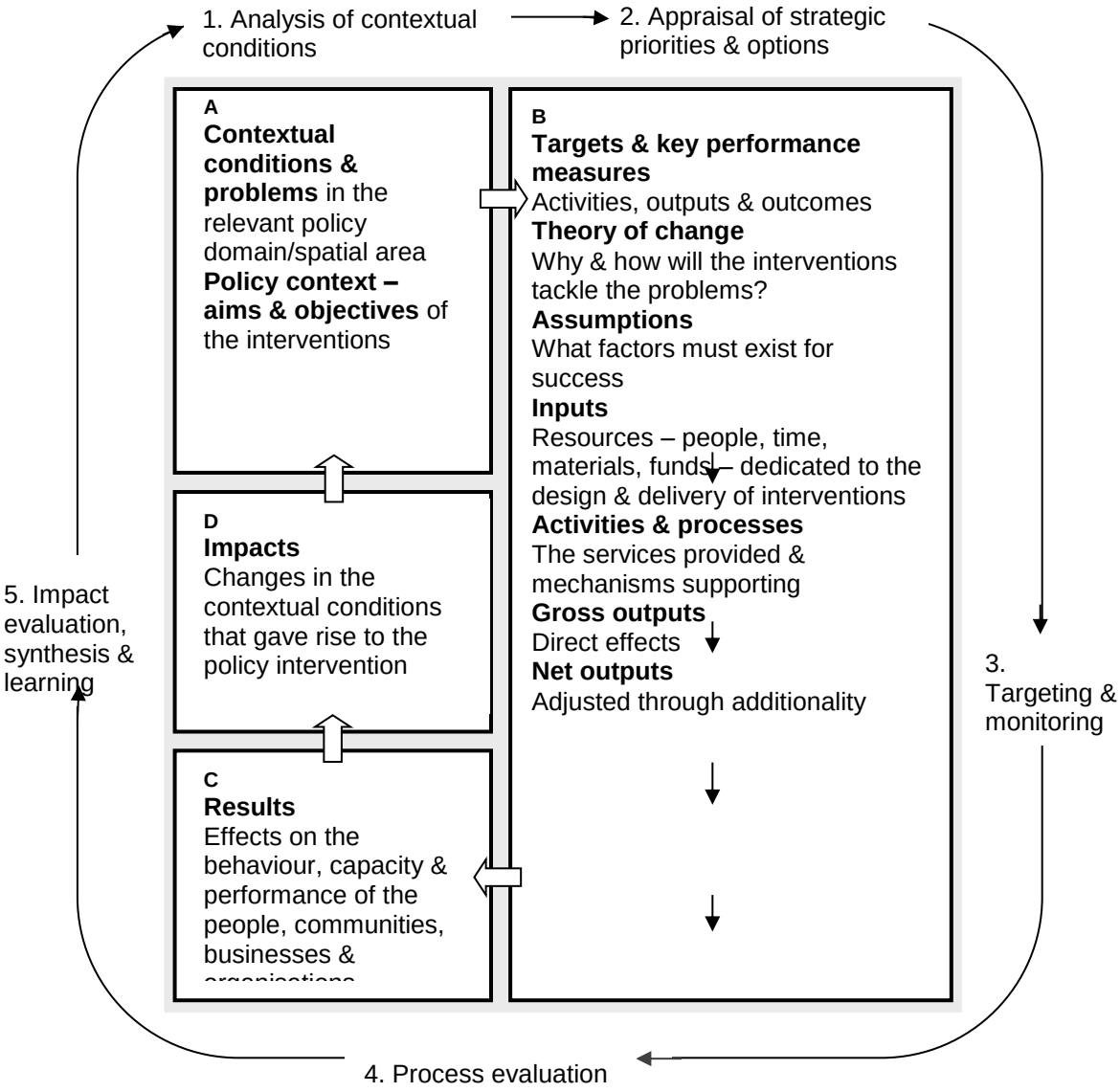
- Do think seriously about resourcing your impact evaluation activity effectively. Make sure that considered justified costs for your impact evaluation are included in your programme budgets and costs from the start.
- Consider carefully the trade-off between using external contractors to undertake the evaluation as opposed to you progressing it internally. External contractors should bring objectivity/detachment, skills/expertise, and knowledge from elsewhere, all of which are important considerations. Internal evaluation brings detailed programme knowledge, familiarity with the underlying issues and processes, and is often the more economical solution undertaking. Don't forget that external evaluation is often viewed as the 'easier' option because the task is being outsourced. But if you do go down this route, make sure you are involved intimately in the evaluation work, and learn from the evaluators and how they go about their work. This way, even with an externally-delivered evaluation, your own evaluation capability is developing and maturing, making you are still more effective programme manager, monitor, and evaluator for the future.
- In planning ahead for your impact evaluations, do make sure that your data sets — within the programme, partners, and beneficiaries — are in place and can be quickly assessed by evaluation activity. One of the most common reasons for poor evaluation results is that the data with which to tell the story of the programme's performance and achievement are missing, or in poor order. Missing data, irrespective of the quality and achievement of the activity undertaken, will almost always compromise what evaluations can say.

Annex 3 Logic Models

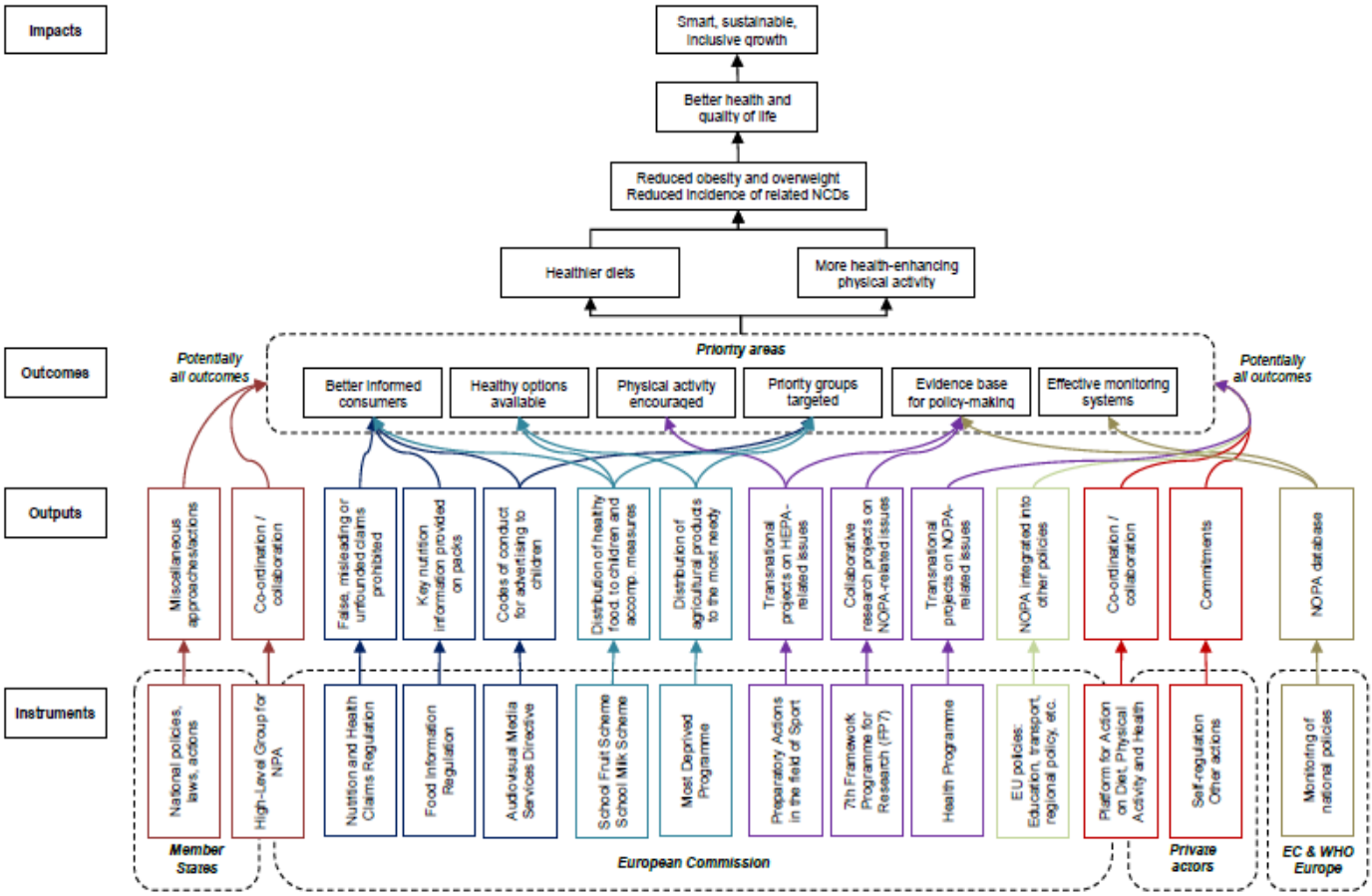
Example: a simple logic model of a business support programme



Example: simple logic model (different format)



Example: a more complicated logic model



Annex 4 Literature

- Howard White Daniel Phillips: Addressing attribution of cause and effect in small n impact evaluations. Towards an integrated framework. June 2012, http://www.3ieimpact.org/media/filer_public/2012/06/29/working_paper_15.pdf

- European Commission: Evalsed Sourcebook – Methods & Techniques

- Rogers, P (2008): Using Programme Theory to Evaluate Complicated & Complex Aspects of Interventions. Evaluation, Vol 14(1): 29 – 48

- Westhorp, G (2014): Realist Impact Evaluation – an Introduction. ODI.org/methodslab

- Mayne, J: Contribution analysis (2008): An approach to exploring cause & effect. ILAC Brief 16

- Baptist, C & Befani, B (2015): Qualitative Comparative Analysis – A Rigorous Qualitative Method for Assessing Impact

- More information on different evaluation approaches can also be found under: http://betterevaluation.org/approach/realist_evaluation