

Evaluating whole programmes? Or delimiting the scope and scale of Evaluations

Elliot Stern

Input into Evaluation Helpdesk Training Seminar

Burgas July 2019

Evaluating whole programmes?

- One of the fundamental design choices that all evaluators face is how to delimit the scope and scale of an evaluation
- As already suggested Evaluation Questions should drive evaluation design – this includes scope and scale
- As also suggested EQs themselves will be shaped by ‘programme characteristics’ – which will depend on administrative and political considerations as well as on the primary purpose of an intervention

Evaluating whole programmes?

The lexicon of policy evaluation usually distinguishes between:

- Strategies
 - Policies
- Programmes
 - Projects
- Actions or activities

The usual assumption is that each level incorporates the one below.....

This is not always so in practice

- We know that it is easier to evaluate the bottom than the top of list
 - Sometimes in evaluation we trade-off ease against relevance!

Evaluating 'whole programmes'?

- There is no obligation to evaluate whole programmes in the CPR but evaluating OPs or other more holistic or cross-cutting units of analysis may sometimes be justified
 - The focus of this session is to identify the conditions under which more extensive units of analysis are justified
- OPs are therefore treated as a subcategory of such a 'more extensive unit of analysis'

Evaluating whole programmes?

There are a number of justifications for a more inclusive unit of analysis

- Necessary to answer a particular EQ
- A relatively narrowly focused programme
- Scale & coverage– e.g. an OP that equates with an entire national, regional or sectoral policy
 - Interdependence and complementarity within a single OP

Evaluating 'whole programmes'?

Answering a particular Evaluation Question

- Some EQs require gathering evidence from across several project spheres within a programme, e.g.

Which strategies for SME support made the greatest contribution to export-led growth?

- In order to answer this EQ in a Sectoral OP at national level, one would need to look across the OP and possibly relate it to national OP priorities (e.g. for export led growth)
- Here the definition of 'whole programme' is those aspects of the whole programme and adjacent programmes that address the EQ

Evaluating whole programmes?

- It may also be necessary to gather evidence across different units – Territorial, Sectoral and National OPs, as well as particular Projects & Policies in order to answer an EQ
- For example we may want to know how transport investments that are sustainable also contribute to labour market policies that are more inclusive
- Here the definition of the ‘whole’ includes all those linked transport interventions, projects and strategies that are judged to influence labour markets inclusivity

A similar logic would apply to the evaluation of Priority Axes or Thematic Areas

Evaluating whole programmes?’

The ‘focus’ of a programme

- Some OPs are narrowly focussed, e.g. on transport infrastructure, urban regeneration or environment
- Other OPs may be national, economy-wide, multi-sectoral and even multi-fund
- It is also more likely that the elements of the narrowly focussed OP are conceived of as interdependent – but not always
- Here we define the holistic unit of analysis pragmatically – that subset of OPs that are relatively well-focussed

Evaluating whole programmes?’

Scale and Coverage

- Some OPs because of their scale and coverage equate to entire policies or national strategies
- It can be argued that such inclusive OPs have to be evaluated holistically because of their ‘global’ coverage

Evaluating whole programmes?

A coherent, multi-measure OP

- The main methodological justification for following a more holistic evaluation strategy is that the elements of a programme are complementary and mutually reinforcing
- An OP wide evaluation would be justified in an OP where the different elements have been designed from the beginning to complement and reinforce each other
 - There may be a similar (or stronger) justification if an overarching policy envisages a set of nested interventions across different sectors and territories – the evaluation of a subset of interventions would not allow for sufficient focus on multi-level linkages and reinforcement

Evaluating whole programmes?

Two notes of warning.....

- First we know that many OPs are not 'programmatic' rather administratively convenient 'baskets' of projects
- Second, we know that evaluation budgets are limited and choices have to be made about resource allocation
- This highlights the importance of a careful pre-analysis of OP content as well as preparing an overall Evaluation Plan before committing to any form of OP-wide evaluation commitment

Evaluating whole programmes?

- The kinds of methods and techniques needed to evaluate more holistic units of analysis will rely on similar families of methods as in other evaluations
 - These would include statistical analyses, cross case or cross beneficiary comparisons, analyses of costs, re-constituting and testing Theories of Change, contribution analysis, longitudinal case studies etc.
- Arguably where there is greater interdependence and complexity there will be a need for more 'complexity appropriate' methods such as econometrics, statistical modelling, systems mapping, process tracing etc.

Evaluating whole programmes?

- Programmes have always been one focus of EU evaluations following one of two routes, top-down and bottom-up
- The bottom-up route starts by evaluating specific actions and projects and aggregates observed effects to confirm the presence of policy priorities
- The top-down route starts with priority themes and searches for examples across programme or multi-programme investments

Evaluating whole programmes?

Following a bottom-up route starting with projects and actions implies:

- Selecting /sampling projects
- Deciding what aspects of these projects should be analysed
- Aggregating project data from selected projects to be able to take a more inclusive, programmatic view

Evaluating whole programmes?

Following the top-down programme route implies

- An analysis of the coherence and design of programmes – often politically sensitive
 - Credible indicators - that relate plausibly to intended results
- The selection of cross-cutting themes, strategies or factors that can be shown to be critical for programme success

Evaluating whole programmes?

Notions of *coherence* at different levels of analysis are central to programme evaluations. For example it is common to consider:

- Internal coherence
- External Coherence and
- Strategic coherence

Evaluating whole programmes?

By internal coherence we mean....

- How well projects and activities/interventions fit together and complement each other, creating positive synergies
- This depends on a theory both about the links in a causal chain and the interdependence between causal factors
- Methodologically internal coherence usually depends on inter-project comparisons and re-constituting Theories of Change

Evaluating whole programmes?

- By external coherence we mean.....
 - How well matched programme activities/interventions are with policy priorities and the needs of intended beneficiaries.
- This depends on a theory of how programme elements lead to policy outcomes and are expected to have positive benefits for beneficiaries
- Methodologically external coherence relies more on both programme and domain theory (i.e. theories about socio-economic behaviour); collating results – or distance travelled - and beneficiary feedback

Evaluating whole programmes?

By strategic coherence we mean....

- How well policies are themselves likely to contribute to overall Strategic priorities, such as Smart Growth; or sustainable development
- Relies more heavily on domain theory – theories of innovation; growth; sustainability etc
- Requires an understanding of how different programmes and policies interact e.g. how does an adjacent or related policy or programme contribute to the focal programme's effectiveness...

Evaluating whole programmes?

A second framework commonly used when evaluating whole programmes is to examine the coherence across the programme's life-cycle:

- Ex ante assessments of needs
 - Programme design
- Implementation, outputs and steering
 - Outcomes and impacts
- To what extent did the programme design meet assessed needs and priorities? Are outcomes and impacts as anticipated at the programme design stage?

Combining coherence frameworks

Coherence	<i>Ex ante needs and starting conditions</i>	<i>Design and partnership set-up</i>	<i>Implementation, outputs & 'steering'</i>	<i>Results and impacts</i>
Internal				
External				
Project				

The diagram consists of a 4x5 grid. The top row contains the following text: **Coherence**, *Ex ante needs and starting conditions*, *Design and partnership set-up*, *Implementation, outputs & 'steering'*, and *Results and impacts*. The three rows below are labeled Internal, External, and Project. A horizontal blue arrow points from the left edge of the top row to the right edge of the top row. A vertical blue arrow points from the bottom edge of the top row to the bottom edge of the bottom row, passing through the middle of the Internal, External, and Project rows.

Evaluating whole programmes?

In reality any single programme evaluation will prioritise some but not all the cells in such a matrix, shaped to a large extent by the operational, planning and learning requirements of Programme stakeholders; and the reporting and accountability requirements set by programme sponsors, in this case national authorities and the European Commission

Such a framework highlights some key interactions.....

- Influence of starting conditions including previous project and policies & evaluations - as precursors for subsequent results and impacts
- Design, including the strength of intervention logic/ToC assumptions; and how partnerships affect the efficiency and effectiveness of programme implementation
- How steering arrangements – a basis for adaptive programming and ‘mid-course’ correction in the 2013-2020 programming period - depends on the strength of partnerships and the active engagement of partners - and their ability to use evaluation outputs

Conclusions

- All evaluations must decide about their scope and scale. Deciding to evaluate 'whole programmes' is part of the same challenge
- It can be easier to evaluate Projects and Interventions rather than Policies and Programmes – may risk trading-off ease with relevance
- No obligation to evaluate OPs and policies therefore need to decide when important to do so
- At centre of most attempts at holistic evaluations are concerns about coherence: within programmes & in relation to external priorities – and across the programming lifecycle

Checklist for Group Session

1. In your OP of interest are there Evaluation Questions that require some kind of 'whole programme' evaluation?
2. What other conditions do you think might justify a more 'holistic' unit of evaluation analysis e.g. whole programmes; cross-cutting programmes; combining OP and Policy evaluation?
3. How practical is it for you to evaluate the internal, external and strategic coherence of OPs?
4. Do you routinely take account of starting conditions and previous evaluations when assessing programme success?
5. Will stakeholders and partners be interested to use programme-wide evaluations as part of mid-term steering and re-programming?