

Q1. Very typically, programme directors are quite resistant to putting in the time and effort to understand the theories that drive their programmes. How do you get the programme directors to take the time and interviews in order to get a proper evaluation?

Suggest to not structure it as a formal interview. Try to meet them at/before/after their project team meetings. Evaluators can utilise background documents to gather a sense of the goals and activities of the programme, utilise literature to help identify causal links, and have draft theory before meeting them. It may be more efficient to validate initial programme theories instead of requiring programme directors to develop theories from scratch. A realist synthesis can bring good inspiration to the conversation, by focusing attention on what we already know and what is new and potentially exciting about the current programme.

Q2. If you are doing an impact evaluation, do you still sit with the programme developers and ask them what was their original intention or mechanism of change that they theorised at the start, or focus mainly on what actually happened over the duration of programme delivery?

Same as above. The theory can be refined based on what happened during implementation. However, it will still be useful to validate with developers/people at stake if focus of evaluation will be of relevance to them.

Q3. What are the typical resources/manpower that are needed for this manner of evaluation?

Resources differ based on the overall aims and scope of the evaluation, and the skills will depend on the type of theory-based evaluation that you pursue.

Q4. In the event there are competing theories, how do you determine which one to adopt?

There are several available papers to refer to:

- a. Birken Sa, et al. T-CaST: an implementation theory comparison and selection tool. *Implement Sci.* 2018 Nov 22;13(1):143. doi: 10.1186/s13012-018-0836-4.
- b. Fontaine G, et al. Advancing the selection of implementation science theories, models, and frameworks: a scoping review and the development of the SELECT-IT meta-framework. *Implement Sci.* 2025 May 28;20(1):24. doi: 10.1186/s13012-025-01436-5.

Q5. How do these theory based evaluations dovetail with QI or innovation approaches?

Theory-based evaluations can add insights on the casual mechanism, which can support the design of innovative approaches with higher potential for success. Quality improvement approaches themselves are also amenable of theory-based evaluation.

Q6. What if I'm pulled into a team halfway through an implementation of a programme—how do I evaluate objectively, when I am unable to randomise, participants are fixed, and many stakeholders still have a black box evaluation mindset?

You can still evaluate objectively mid-way through implementation but perhaps not comprehensively as some data may not have been collected from the beginning. It is important to discuss and manage stakeholders expectations of what can be answered given the limitations.

Q7. Will Grounded Theory methods be most appropriate for qualitative explorations?

No necessarily, it depends on your questions. There is a kinship between realist evaluation and its focus on contexts, mechanisms and outcomes, and grounded theory, which focuses on social process. Some flavours of grounded theory use explanatory matrices, which include context, conditions, process, consequences and outcomes, which has a clear affinity with the CMO heuristic used in realist evaluation. However, this link is controversial in some realist circles. In any event, principles of grounded theory, like constant comparison and purposive sampling, can be usefully adapted to your qualitative data collection if your evaluation involves collecting primary qualitative data.

Q8. What are the main practical challenges that you have faced when carrying out theory based programme evaluations, other than those already mentioned?

Programmes are dynamic and it can be challenging for the evaluation team to keep track of changes made to the intervention components and its implementation approach. This however impacts the implementation of the evaluation and interpretation of the evaluation findings.

Q9. To the panellists, what are the typical barriers that you face when doing evaluations? What are the things you wish programme directors would have when they first talk to you?

Common barriers:

- Lack of understanding among stakeholders about evaluation' approaches
- Difficulty in accessing data
- Coming into the conversation too late, when evaluation data collection plans could have been better customised
- Understanding different stakeholders' perspectives of intervention functioning

Evaluators generally find it useful when programme directors have reflected on the 'why' of their programme. Developing programme theories is a partnership activity, but an early intuition is always valuable.

Q10. Question for Mila: Are there ways to properly reverse engineer the theories effectively?

We can take the same approach as building a programme theory for the start of programme. However we maybe limited to available information/data for building/refining of theories.

Q11. Question for Mila: As implementing a new clinical programme is an iterative process, do you generate the theory of change at the beginning or at the end of evaluation?

I usually co-develop ToC before evaluation to help focus on data collection and analysis. Based on the results of evaluation or as programme develops, we adapt ToC accordingly.

Q12. Practically, how do you manage the politics of evaluation, especially when the formative or summative results are poor?

It is important to engage the relevant stakeholders at various stages of the evaluation and to be transparent about the purpose of the evaluation, provide interim findings and to seek their responses, and to communicate the summative findings to seek a common understanding before submission of the findings to the funders. At its most useful, evaluation is a learning activity. Because of this, thinking at the point of evaluation design about how findings can be mobilised can reduce the 'pain' of a poor result.

Q13. Would it be better to have a third party evaluator or to build internal capability if you are in an organisation?

Not all programmes are complex with heavy investment. For complex programmes with significant scope for future replication, a third-party evaluator who may be able to assemble a team with the necessary skills and objectivity might be more appropriate. For programmes whose primary purpose is to improve current care processes, it would make sense to utilise internal programme who are embedded within the system.