Simple but not simplistic: Findings from a theory-driven retrospective evaluation of a small projects program

Description

Tweet

By Larry Dershem, Maya Komakhidze, Mariam Berianidze, in Evaluation and Program Planning 97 (2023) 102267. A link to the article, which will be active for 30 days. After that, contact the authors.

Why I like this evaluation â?? see below and the lesson I may have learned

Background and purpose: From 2010â??2019, the United States Peace Corps Volunteers in Georgia implemented 270 small projects as part of the US Peace Corps/Georgia Small Projects Assistance (SPA) Program. In early 2020, the US Peace Corps/Georgia office commissioned a retrospective evaluation of these projects. The key evaluation questions were: 1) To what degree were SPA Program projects successful in achieving the SPA Program objectives over the ten years, 2) To what extent can the achieved outcomes be attributed to the SPA Program â??s interventions, and 3) How can the SPA Program be improved to increase likelihood of success of future projects.

Methods: Three theory-driven methods were used to answer the evaluation questions. First, a performance rubric was collaboratively developed with SPA Program staff to clearly identify which small projects had achieved intended outcomes and satisfied the SPA Program â??s criteria for successful projects. Second, qualitative comparative analysis was used to understand the conditions that led to successful and unsuccessful projects and obtain a causal package of conditions that was conducive to a successful outcome. Third, causal process tracing was used to unpack how and why the conjunction of conditions identified through qualitative comparative analysis were sufficient for a successful outcome.

Findings: Based on the performance rubric, thirty-one percent (82) of small projects were categorized as successful. Using Boolean minimization of a truth table based on cross case analysis of successful projects, a causal package of five conditions was sufficient to produce the likelihood of a successful outcome. Of the five conditions in the causal package, the productive relationship of two conditions was sequential whereas for the remaining three conditions it was simultaneous. Distinctive characteristics explained the remaining successful projects that had only several of the five conditions present from the causal package. A causal package, comprised of the conjunction of two conditions, was sufficient to produce the likelihood of an unsuccessful project. Conclusions: Despite having modest grant amounts, short implementation periods, and a relatively straightforward intervention logic, success in the SPA Program was uncommon over the ten years because a complex combination of conditions was necessary to achieve success. In contrast, project failure was more frequent and uncomplicated. However, by focusing on the causal package of five conditions during project design and implementation, the success of small projects can be increased.

Why I like this paper:

1. The clear explanation of the basic QCA process
2. The detailed connection made between the conditions being investigated and the background theory of change about the projects being analysed.
3. The section on causal process which investigates alternative sequencing of conditions
4. The within case descriptions of modal cases (true positives) and the cases which were successful but not covered by the intermediate solution (false negatives), and the contextual background given for each of the conditions you are investigating.
5. The investigation of the causes of the absence of the outcome, all too often not given sufficient attention in other studies/evaluation
6. The points made in the summary especially about the possibility of causal configurations changing over time, and a proposal to include characteristics of the intermediate solution into the project proposal screening stage. It has bugged me for a long time how little attention is given to the theory embodied into project proposal screening processes, let alone testing details of these assessments against subsequent outcomes. I know the authors were not proposing this specifically here but the idea of revising the selection process by new evidence of prior performance is consistent and makes a lot of sense
7. The fact that the data set is part of the paper and open to reanalysis by others (see below)

New lessons, at least for me..about satisficing versus optimising

It could be argued that the search for Sufficient conditions (individual or configurations of) is a minimalist ambition, a form of â??satisficingâ?• rather than optimising. In the above authorsâ?? analysis their â??intermediate solutionâ?•, which met the criteria of sufficiency, accounted for 5 of the 12 cases where the expected outcome was present.

A more ambitious and optimising approach would be to seek maximum classification accuracy (=(TP+TN)/(TP+FP+FN+TN)), even if this at the initial cost of few False Positives. In my investigation of the same data set there was a single condition that was not sufficient, yet accounted for 9 of the same 12 cases (NEED). This was at the cost of some inconsistency i.e two false positives also being present when this single condition was present (Cases 10 & 25) . This solution covered 75% of the cases with expected outcomes, versus 42% with the satisficing solution.

What might need to be taken into account when considering this choice of whether to prefer optimising over satisficing? One factor to consider is the nature of the performance of the two false positive cases? Was it near the boundary of what would be seen as successful performance i.e. a near miss? Or was it a really bad fail? Secondly, if it was a really bad fail, in terms of degree of failure, how significant was that for the lives of the people involved? How damaging was it? Thirdly, how avoidable was that failure? In the future is there a clear way in which these types of failure could be avoided, or not?

This argument relates to a point I have made on many occasions elsewhere. Different situations require different concerns about the nature of failure. An investor in the stock market can afford a high proportion of false positives in their predictions, so long as their classification accuracy is above 50% and they have plenty of time available. In the longer term they will be able to recover their losses and make a profit. But a brain surgeon can afford absolute minimum of false positives. If their patients die as a response of their wrong interpretation of what is needed that life is unrecoverable, and no amount of subsequent successful future operations will make a difference. At the very most, they will have learnt how to avoid such catastrophic mistakes in the future.

So my argument here is letâ??s not be too satisfied with satisficing solutions. Letâ??s make sure that we have at the very least always tried to find the optimal solution (defined in terms of highest classification accuracy) and then looked closely at the extent to which that optimal solution can be afforded.

PS 1: Where there are â??imbalanced classesâ?• i.e a high proportion of outcome-absent cases (or vice versa) an alternate measure known as â??balanced accuracyâ?• is preferred. Balanced accuracy = (TP/(TP+FN))+(TN/(TN+FP)))/2.

PS 2: If you have any examples of QCA studies that have compared sufficient solutions with non-sufficient but more (classification) accurate solutions, please let me know. They may be more common than I am assuming

Category

1. Journal article
2. Process tracing
3. Qualitative methods

Date
13/09/2026
Date Created
13/03/2023
Author
admin