



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

Larry Dershem^{a,*},¹, Maya Komakhidze^b, Mariam Berianidze^c

^a Policy and Management Consulting Research Center, Tbilisi, Georgia

^b Lynch School of Education and Human Development, Boston College, Massachusetts, USA

^c Policy and Management Consulting Research Center in Tbilisi, Georgia

ARTICLE INFO

Keywords:

Intervention logic
Theory-driven evaluation
Participatory approach
Performance rubric
Multi-method
Qualitative comparative analysis
Causal process tracing

ABSTRACT

Background and purpose: From 2010–2019, 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'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'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.

Johan Cruyff - "Soccer is simple, but it is difficult to play simple."

1. Introduction

In March 2020, the United States Peace Corps in (the republic of)

Georgia commissioned a ten-year (2010 – 2019) retrospective evaluation of their nation-wide, local development-focused Small Projects Assistance (SPA) Program. During these ten years, 270 small projects were implemented overwhelmingly by young (mid-twenties) Peace Corps Volunteers (PCVs) as a secondary project to their primary

* Correspondence to: PMCG, 61 Davit Aghmashenebeli Ave, Tbilisi 0102, Georgia.

E-mail addresses: l.dershem@pmcginternational.com (L. Dershem), mayakomak@gmail.com (M. Komakhidze), m.berianidze@pmcginternational.com (M. Berianidze).

¹ ORCID: 0000-0002-7603-793X

<https://doi.org/10.1016/j.evalprogplan.2023.102267>

Received 9 July 2022; Received in revised form 7 February 2023; Accepted 24 February 2023

Available online 28 February 2023

0149-7189/© 2023 Elsevier Ltd. All rights reserved.

assignment. Due the covid-19 pandemic in Georgia, in early 2020, Peace Corps Georgia (PC/Georgia) evacuated all PCVs from Georgia to the US with the expectation of PCVs returning to Georgia once the covid-19 infection rate was low and manageable. Before new PCVs returned to Georgia and got involved in the SPA Program, PC/Georgia commissioned a retrospective evaluation to answer a set of key evaluation questions (KEQs) to help improve SPA Program's guidance documents, the grant review process, and training for PCVs before implementing a small project: 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's interventions,* and 3) *How can the SPA Program be improved to increase likelihood of success of future projects.*

This article aims to show how a theory-driven, mixed method approach is used to answer these three KEQs in a retrospective evaluation. First, we present the development of a performance rubric to determine successful and unsuccessful projects among the 270 small projects conducted over 10 years. Next, we present the use of qualitative comparative analysis (QCA) for case comparison based on logical procedures (Boolean algebra) that minimize the configuration of necessary and sufficient conditions that distinguish between successful and unsuccessful small projects. Last, we present the use of causal process tracing, a within-case observational analysis, to unpack the causal and temporal processes behind the QCA's intermediate solution for successful small projects.

According to Rodgers (2008) and (Coryn, et al., 2011), theory-driven evaluation provides one of the most advantageous approaches to answering these types of KEQs since it focuses evaluation findings on whether a program has been effective, illuminating the program's underlying causal mechanisms to achieve planned objectives, and identifying areas for improvement for replication of success. The authors use a theory-driven approach in this evaluation because it prioritizes methodological pluralism and evidential diversity in addressing KEQs in contrast to approaches that prioritize a method (i.e., quasi/experimental designs), as exemplified by What Works Clearinghouse (2008) and Campbell Collaboration (2019). Theory-driven approaches are method neutral (Pawson & Tilley, 1997).

Theory-driven evaluation has two essential components (Coryn, et al., 2011), conceptual and empirical. The conceptual component involves clarifying the program's theory or model, however, as argued by Hawkins (2020), the conceptual component should focus not on a program's theory because "...programs are not, on close examination, best understood as theories at all....The core essence of a program is considered to be a proposition about the value of a particular course of action: a plan or argument drawing on a set of reasons to suggest it will be effective," in other words, the conceptual component should focus on a program's intervention logic used for a specific purpose in a specific context.

Conceptualizing a program as a specific intervention logic or a particular course of action, rather than a theory, makes it easier to understand and evaluate logically, which leads to the empirical component of theory-driven evaluation. The empirical component of theory-driven evaluation seeks to investigate how the program caused the intended outcomes, but is method-neutral or methodologically pluralistic, thus, there is no "gold standard" since no method has primacy in answering these evaluation questions (Donaldson, 2007; Chen, 2005).

A program's intervention logic is typically represented in a graphic that depicts relationships among events, such as activities, outputs, short-term outcomes, long-term outcomes/impacts, and other factors, although they also may be expressed in tabular, narrative, or other forms (Rodgers, 2008; Coryn, et al., 2011). Because of the limited number of events (inputs, activities, outputs, outcomes, impacts) and depiction of direct, event-based causality (e.g., discrete occurrences within a delineated time that are characterized by a beginning and an ending), these logical models are often described as simple, linear program models. As described by Rodgers (2008), simple program logics represent "following a cooking recipe," complicated program logics represent "sending a rocket

to the moon," and complex program logics represent "raising a child." However, although a program's intervention logic may appear to be simple, it does not warrant the assumption it is at best, simplistic, or at worst, incorrect. Rather, the level of a causal explanation should be sufficient to the explanatory task at hand (Grotzer & Perkins, 2000). A sufficient causal explanation to a student driver of how to stop a moving car is simply and sufficiently "push down on the brake pedal" rather than explaining the complex relationships between the brake pedal, master cylinder, hydraulic fluid and pressure system, brake fluid pipes, brake calipers, brake pads, wheel cylinders, and so forth.

As contended by Hawkins (2020), all the components depicted in an intervention logic, whether simple, complicated, or complex, should be viewed as a "causal package" and not a "causal chain." That is, the belief by program designers is that the program only as a whole is sufficient to bring about intended outcomes and that each component of the intervention logic, although needed, is not sufficient to bring about the intended outcomes. During the implementation of the project, the causal package is comprised of complex patterns of interaction and relationships among people and groups involved in activities and events that occur in, and are influenced by, supporting and disrupting conditions in particular context (e.g., history, culture, power relations) to achieve intended short-term and long-term outcomes.

Invoking necessary and sufficient conditions to achieve intended outcomes, even in a simple program design logic, suggest INUS causality (Mackie, 1974). INUS causality refers to a cause as an Insufficient (i.e., inadequate alone) but Nonredundant (i.e., though still necessary) part of an Unnecessary (i.e., there are other ways) but Sufficient (i.e., adequate) condition. An example of INUS causality is that a lit cigarette alone is insufficient to start a large fire, though a necessary ignition source, but requires other conditions to start a large fire, such as oxygen, adequate fuel (e.g., dry grass, trees), and strong winds all of which are necessary but none of them alone are sufficient to cause a large fire; rather, it is the combination, or configuration, between these necessary conditions that make them sufficient to produce the outcome, a large fire. And there are other ways large fires (i.e., outcomes) can occur depending on the context. Similarly, in a simple program design logic, inputs, activities, and outputs alone and in isolation do not lead to outcomes but rather the proper interaction and configuration of them at the right time and context, generating "productive continuity" (Machamer et al., 2000) that can make a meaningful difference in achieving intended outcomes.

2. The evaluand

This retrospective evaluation of the SPA Program in Georgia is the first evaluation of the SPA Program in Georgia. The only prior evaluation of the Peace Corps' SPA program covered Morocco, Philippines, Rwanda, Peru, and Malawi from 2013 to 2017 (Muco et al., 2018) and given its international scope, the recommendations produced were global in nature and less tailored to the specific national contexts where Peace Corps' SPA Program operate. This retrospective evaluation of the SPA Program in Georgia was designed to be locally focused and informative for planning and implementing an effective small grants program within the developmental and cultural context of Georgia.

As mentioned earlier, the US PC/Georgia SPA Program is implemented by PCVs and this retrospective evaluation covered small projects implemented from 2010 to 2019. The focus of the retrospective evaluation was to provide the SPA Program with a sense of how successful small projects had accomplished program objectives over the ten years, understand what components of their intervention logic were crucial in leading to success of small projects, and how could the implementation of small projects be improved.

These are called small projects because the maximum budget was \$2500 USD and PCVs who wanted to supplement their primary assignment with a small project could obtain a grant six months after being posted with the requirement the project be completed three months prior to the completion of their two years of service; thus, the maximum

timeframe for a project was 15 months although there are a few exceptions. The two primary objectives of the SPA Program (United States Peace Corps, 2015) were: 1) *capacity building of community members to identify and prioritize what they would like to change and to use their own strengths as well as learn new skills to achieve that change* and 2) *sustainability or the ability of a project to continue to meet the needs of a community once the initial SPA Program grant has ended*. Other requirements of these projects were:

- Community initiated, designed, and directed to serve the needs of the community where the PCV is posted.
- The PCV should partner with a local counterpart to help and support in the design and implementation.
- Involves capacity-building through developing or strengthening the skills, competencies, and abilities of individuals, organizations, or institutions in the community.
- Ensures the community contributes, at least, 25% of the total grant budget, which can be in cash or in-kind contributions.
- Involves activities the local community or organization can continue (or expand) on its own, following the PCV's close of service.
- Includes monitoring and evaluation to measure and document results.
- Must be submitted and approved by the SPA Review Committee comprised of PC/Georgia staff and PCVs.

All PCVs involved in the SPA program during these years were recent college graduates with a bachelor's degree (a PC requirement) who ranged from 22 to 26 years of age with little to no experience in local development and project management. From 2010–2019, a total of the 270 small projects were implemented by these young PCVs throughout Georgia. Almost one-half (47% or 127) were focused to *Education*, such as establishing English language classrooms in schools, restoration of libraries, providing computer technology and trainings for teachers and students, creating after-school English clubs, planting school gardens, and establishing science labs in schools. Next, in rank order based on number of projects: *Youth Development* (67 or 25%) projects that included conducting summer camps, fitness training & sports, STEM training, and entrepreneurship; *Community Economic Development* (34 or 13%) projects that involved training in small business and financial management, marketing and accounting skills, agritourism development, cottage industry such as craft production, local agriculture production (e.g., honey), and technology use (computers, digital camera); *Health* (16 or 6%) projects focusing on fitness training, healthy lifestyles training, health information, and hygiene; *Environment* (10 or 4%) projects that include eco-education programs, Green camps, and nature conservancy; *Information, Communication and Technology* (5 or 2%) projects involving computer skills, social media, technology camp, and use of new technology in classrooms; *NGO Development* (5 or 2%) projects focused on organizational capacity building, enhancing visibility, fund-raising, development of resource & training room; *Gender Issues* (2 or 1%) projects that provided training for new mothers, self-esteem & leadership, and health awareness; *Volunteerism* (2 or 1%) projects connecting and coordinating volunteers and capacity building for volunteers; a *Municipal Development* (1 or 0.5%) project that provided skills training for staff of a youth house; and a *Water & Sanitation* (1 or 0.5%) project that involved construction of toilets in a school and raising hygiene awareness. Most (78% or 210) of the 270 SPA Program projects were implemented in small towns and villages while all of the remaining projects were implemented in large municipalities.

2.1. Data collection methods and sources

As the small projects included in the evaluation were implemented in the period between 2010 and 2019, the PCVs who had implemented projects had already completed their assignments and left Georgia.

Additionally, because of the social distancing measures introduced in Georgia, researchers' ability to visit the communities where the projects were implemented was limited. Consequently, the evidence for this evaluation was primarily obtained from online interviews and focus group discussions as well as an extensive desk review. Desk review entailed examination of project documents in the SPA Program database, which included small project guidance documents, PCV Small Project Handbook, small project grant applications with details of required project description, revised small project designs after incorporating the review committee's comments, monitoring reports, and the small project completion reports. Project design, monitoring and completion reports had consistent structure which aided in the collection of standardized data across projects.

Additional data and evidence were obtained from online data collection with: 1) in-depth interviews with four current senior SPA Program staff, one with knowledge of small projects going back more than eight years, 2) one online focus group discussions with five PCVs who were members of the 2019 SPA Program Review Committee, 3) one in-depth interview with a representative from the funding agency (USAID), 4) an online survey of 56 PCVs, using Google Forms, who implemented a small project during the decade being reviewed, 5) in-depth interviews with 25 PCVs who had primarily served in the last five years, 6) in-depth interviews with 35 Counterparts of PCVs who assisted with the small project, and 7) a mixture of online/phone interviews and FGD with 57 direct beneficiaries from 35 small projects implemented in the last 5 years.²

2.2. Theory-driven, multi-method evaluation approach

As mentioned earlier, the KEQs for this retrospective evaluation were: 1) *To what degree were SPA Program projects successful in achieving the SPA Program objectives?*, 2) *To what extent can the achieved outcomes be attributed to the SPA Program's interventions?*, and 3) *How can the SPA Program be improved to increase likelihood of success of future projects?* and that theory-driven evaluation approach was most suitable to answer these evaluation questions. In keeping with Rodgers (2008), theory-driven evaluation "refers to a variety of ways of developing a causal model linking programme inputs and activities to a chain of intended or observed outcomes, and then using this model to guide the evaluation."

To help answer the first KEQ, regarding success in achieving program objectives, a performance rubric was developed with current SPA Program staff to clearly articulate criteria and definitions of performance in achieving the intended outcomes (King et al., 2013). Or in other words, the performance rubric helped in the evaluative judgement of determining "how good is good" and "how good is good enough" (Davidson et al., 2011) in the eyes of those who commissioned the evaluation.

To help answer the second KEQ, regarding attributing achieved outcomes to project interventions, two approaches were used. The first approach involved developing a graphic depiction of the SPA Program's intervention logic, in collaboration with SPA Program staff and based on

² Primary data collection among community members, PCV counterparts, and PCVs was severely constrained in this evaluation for multiple reasons. First, small project documentation kept records of the number of beneficiaries but no contact information, which made identifying and locating beneficiaries difficult. The beneficiaries interviewed in this evaluation were identified by the assistance of PCV counterparts who recalled some of the community members who participated in the project. Second, the evaluation was commissioned in March 2020 at the start of covid-19 pandemic, during the duration of this evaluation travel and in-person meetings were restricted and, in some cases, prohibited; thus, interviewing community beneficiaries and PCV counterparts occurred on the phone or online. Third, due to privacy restrictions, the PC Georgia office did not provide the evaluators with contact information of previous or current PCVs to the evaluators. Instead, the PC Georgia office sent an email to all PCVs in Georgia from 2010 to 2019 they had on record requesting them to participate in this evaluation.

various program documents, to guide the evaluation. The second approach involved the use of QCA to identify causal packages or configurations of necessary and sufficient conditions for project success.

To help answer third KEQ concerning program improvement, following Blatter and Haverland (2014), we used causal process tracing (CPT) since QCA only identifies necessary and sufficient causal packages across cases, but it does not provide an explanation of the connections between conditions comprising the causal package. In contrast to process tracing, CPT starts with a proposed causal package then returns to the cases that are members of this causal package to “unpack” the causal mechanisms behind this pattern (Illari, 2011; Russo & Williamson, 2007). Causal process tracing a Theory of Change, or in the case of this evaluation, an intervention logic, can be very helpful for program prediction, planning, and evaluation (Cartwright et al., 2020).

2.3. Constructing the SPA program's intervention logic

The first step to answer the KEQs was to understand the program's intervention logic, since it is fundamentally the unit of analysis in a theory-driven approach (Dalkin et al., 2015), which meant diagramming the set of critical components of the SPA Program and how each component was logically related to each other to achieve desired outcomes (Hawkins, 2020) to help guide the retrospective evaluation. Even though the SPA Program's intervention logic was based on various theories of change (i.e., community development, community capacity building, community psychology, adult learning, and social capital), nonetheless, the evaluation focused on the SPA Program's chosen course of action toward local development that could be empirically explored and assessed; that is, its intervention logic. Assisted by the evaluation team, the SPA Program staff constructed a framework of the program's intervention logic based on SPA Program policy and guidance documents as well as practices over the ten years. Fig. 1 presents the SPA Program's intervention logic framework.

The SPA Program's intervention logic framework starts with the PCV and the PCV's local “community counterpart” working together on Community Identification and a Needs Assessment. Community Identification entailed clearly determining who in the community should be involved in and directly benefit from the project. Once the direct beneficiaries were identified, then a systematic and thorough Needs Assessment should be conducted. Next the Project Design training involved the PCV and their Community Counterpart attending a project design workshop facilitated by SPA Program staff and afterwards the SPA Program review committee reviewing the project design to ensure key inputs and activities are incorporated.

Inputs include all the resources to complete the activities, such as grant funding, community contribution, procuring supplies and materials, and identification of capacity building trainers. Activities include the accomplishment of training, workshops, and occasionally infrastructure rehabilitation. Three primary intended Outputs are the number of community members participating in the project, the number of community members participating in capacity-building trainings, and number of new or improved infrastructure, equipment, and resources. There are four intended Outcomes: a) community members with increased capacities, b) capacities of local organizations strengthened, c) improved local infrastructure, and d) increased well-being of community. All project outcomes were to be accomplished three months prior to the end of the two-year posting of the PCV. The longer-term Impacts of sustained capacity of local communities to conduct low-cost grass-roots development, a more informed engaged citizenry, and more resilient communities were expected to occur at some point in time after the SPA project ended. However, with the PCV's service ending the SPA Program did not fund any follow-up monitoring to assess if these impacts were achieved.

2.4. Creating a performance rubric to identify (Un)successful small projects

After developing the intervention logic depicting the causally related sequence of interventions leading to intended outcomes, the evaluation team worked with SPA Program staff to determine what a successful small project looked like to answer the first KEQ, “To what degree were SPA Program projects successful in achieving the SPA Program objectives?” To determine if a small project had been successful or not, the evaluation team, with the SPA Program staff and the SPA Program evaluation committee, developed a performance rubric, which is a shared understanding articulated in descriptors as well as criteria of what different levels of performance looks like (King et al., 2013). The four descriptors of performance to clarify “how good is good” and “how good is good enough” were: 1) Poor (i.e., not good enough) - representing unsatisfactory performance, 2) Adequate (i.e., not good enough) - representing a fair performance overall with some troublesome weaknesses, 3) Good (i.e., good enough)- representing a satisfactory performance overall though with a few slight weaknesses but nothing of real consequence, and 4) Excellent (i.e., good) - representing an exemplary performance or very good practices. The Project Performance Rubric, shown in Table 1, presents the criteria and standards for each of the four project performance categories.

The performance rubric was designed to give operationalizability to each category and simultaneously offer sufficient flexibility to make the categories qualitatively meaningful. Although a project might score well on criteria that appear in more than one of the four categories, a project was categorized primarily based on the “Necessary” criteria identified in the rubric, which prioritized sustainability and community ownership. Initially, a project was placed in the highest category for which it satisfied the “Necessary” criteria and the project was kept in that category if it met at least two of the additional criteria. Conversely, if the project failed to satisfy any of the additional criteria, the project was downgraded to the next lowest category, and the project was then reassessed based on the criteria of the lower category.

Using the performance rubric, the evaluation team reviewed all project-related documents, such as the application form, project monitoring records, and the project completion report, for each of the 270 small projects and placed each small project into one of the four performance categories. After all projects had been categorized, the results were presented to the SPA Program staff for review. The evaluation team sought SPA program staff's feedback on the performance categorization of each small project and any differences were reconciled. Because some of the projects, especially for those projects conducted eight-to-ten-year prior, had insufficient documentation, if several critical criteria were not met (P1 and P2, or P1 and P3, or P2 and P3) they were classified as Poor. For a project to be classified as Good or Excellent enough evidence, via documentation, was necessary.

After reviewing all 270 projects according to the performance rubric criteria, 21 (8%) were rated Excellent, 61 (23%) were rated Good, 157 (58.1%) were rated Adequate, and 31 (12%) were rated Poor. The categories “Excellent” and “Good” were determined to be “good enough” or “Successful” small projects and the categories of Poor and Adequate were determined to be “not good enough” or “Unsuccessful” small projects. Fig. 2 presents the number of projects rated Successful and Unsuccessful over the ten years and overall.

Based on the performance rubric, the answer to the first KEQ is that one out of every three (30.3%) SPA Program projects over the ten years produced desired and intended outcomes of the SPA Program. In discussion with SPA Program staff, this result did not surprise them, which is one reason the SPA Program commissioned this evaluation. Next, the SPA Program staff wanted to know what components of the program's intervention logic and/or other conditions contributed to a successful small project.

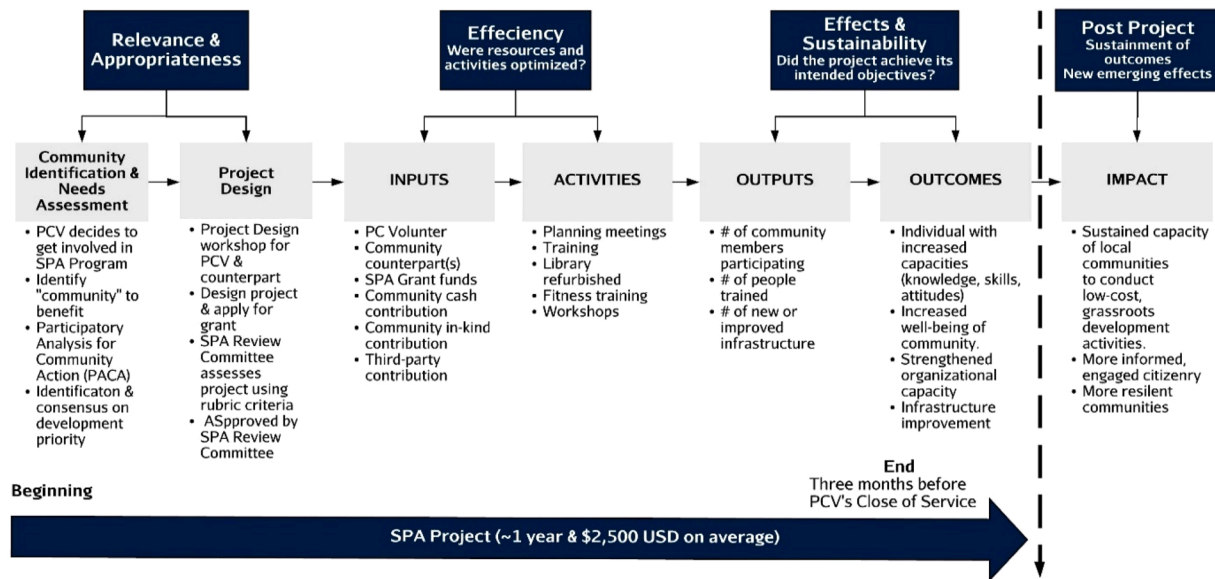


Fig. 1. SPA Program Intervention Logic Framework.

2.5. Cross case comparison of Un/successful small projects

To answer the second KEQ, of what parts of the intervention logic could be causally attributed to a successful project, QCA was used. QCA is one of several set-theoretic methods with the possibility to identify conditions, or configurations of conditions, that are necessary and sufficient for an outcome. Theoretic sets represent well-defined groups where cases comprising the groups are members of the set. In this evaluation, cases are the SPA Program small projects. A crisp-set QCA was used, which involved assigning cases to the presence (=1) or absence (=0) of a condition.³ For the Outcome, set membership was either 1 =Successful or 0 =Unsuccessful small project based on the performance rubric. For the conditions, which from the intervention logic are expected to be causally related to the outcome (Rihoux & Ragin, 2009), SPA Program staff and project stakeholders identified 40 possible conditions they thought were causally related to a successful outcome whether components of the intervention logic or not.

QCA not only examines the causal relation of each condition to the outcome, but also all the possible combinations of the presence and absence of each condition and the number of logically possible combinations of conditions exponentially increases with the number of conditions. Since each condition can be present or absent in each case, the formula for calculating the number of possible logical combinations is 2^k , with k being the number of conditions; thus, 3 conditions result in 8 possible combinations, 4 conditions result in 16 possible combinations, and 10 conditions result in 1024 possible combinations. Following the best practices of QCA approach, the list of 40 conditions was reduced to a list of 10 conditions that met the following requirements: 1) *repetition*, most stakeholders identified the condition, 2) *operationalizability*, the condition could be determined to be present or absent based on project documents and/or interviews, and 3) *variance*, the presence or absence of the condition varied sufficiently across the cases. The result was the final shortlist of 10 conditions that were believed to be related to a successful project.

The 10 conditions that met all three requirements for inclusion in the QCA were:

1. Needs Assessment Conducted – Whether a project was based on a sufficiently well-conducted needs assessment or primarily relied on a PCV's and/or counterpart's preconceived notions about a suitable community project.
2. Beneficiary Age Group – Whether a project primarily targeted youth or had an older population as a beneficiary target group.
3. Length of Project – Whether the project implementation period was shorter or longer than 6 months.
4. PCV in the first or second year of assignment – Whether the PCV who was the primary implementer of the project was in the first or second year of his/her assignment.
5. Capacity-building was core objective of project – Whether the main component of the project revolved around capacity building or primarily involved the purchase/repair of infrastructure and equipment.
6. Who conducted the capacity-building training – Whether capacity-building activities within the project were delivered by the PCV or local counterpart.
7. PCV received administrative and logistical support – Whether the PCV had reliable and persistent support from a local counterpart in dealing with administrative and logistical aspects of the project.
8. PCV received support from an influential community member – Whether the SPA project had the backing of an influential local community member. This could have also been the counterpart of the project.
9. Project had strategies and activities to increase beneficiary interest and engagement – Whether a PCV had planned specific activities and incentives to increase engagement from community members or conversely, assumed community interest in the project.
10. Post capacity-building hands-on activities – Whether the PCV had planned specific hands-on activities as a follow-up to capacity building or conversely, the project culminated with capacity building.

Similar to the OUTCOME crisp-set, each of the ten conditions were calibrated into a crisp-set indicating either the presence (1) or absence (0) of the condition. Table 2 below presents each of the ten conditions, its location within the intervention logic, the acronym used in the QCA for the condition, whether the condition is present in the Logic Model, and condition's calibration into a crisp-set. For example, the condition,

³ Crisp-set was used because many project documents, going back ten years, did not allow for a finer grain of calibration needed for fuzzy-set QCA.

Table 1
SPA Program Small Project Performance Rubric.

Performance Category	Criteria & Threshold	Outcome Classification
Poor - Clearly shows unsatisfactory functioning or performance	P1. Terminated early or not completed. (Excluding unpreventable causes) P2. Less than half of the objectives achieved. (Necessary) P3. No actual beneficiaries or substantially* less than the # of intended beneficiaries. P4. Sustainability* * - Neither economic nor managerial sustainability was considered.	Unsuccessful
Adequate - Clearly shows fair performance overall however has some troublesome weaknesses	A1. Completed with some delay. (Excluding unpreventable causes) A2. At least half or more objectives were achieved. A3. # of actual beneficiaries was not substantially less than # of intended beneficiaries. A4. Sustainability – Either economic or managerial sustainability was considered. (Necessary)	
Good - Clearly shows good performance overall; might have a few slight weaknesses on PC criteria but nothing of real consequence	G1. Completed on time. (or approximately on time) G2. All objectives were achieved (unachievable objectives were modified/adapted). G3. # of actual beneficiaries same or similar to # of intended beneficiaries. G4. All four dimensions of sustainability were considered. (Necessary) G5. Partial transition of ownership to the community. (Necessary)	Successful
Excellent - Clearly shows exemplary performance or very good practices according to PC criteria	E1. Completed on time (approximately). E2. All objectives were achieved, and the project went beyond them towards the goal. E3. # of actual beneficiaries exceeded the # of intended beneficiaries. E4. All four dimensions of sustainability were considered. (Necessary) E5. Project was scaled up and there is a clear and strong indication that further activities were (older projects) or will be (newer projects) implemented beyond the timeframe of the project. (Necessary in the case of only partial ownership – E6) E6. Partial or full transition of ownership to the community. (Necessary)	

* Substantially- 20% or more.

** Sustainability – 1) **Culturally sustainable** (does the basic approach or concept fit within and build on local beliefs and traditions, or will it be seen as an “outsider’s idea” and not be acceptable or continued when the PCV leaves?), 2) **Politically sustainable** (when there is no longer an outsider, such as a PCV in the project, will it be sustainable within the sociopolitical context?), 3) **Economically sustainable** (will there be sufficient local resources or the capacity to generate them when supportive outsiders, such as PCV leaves?), and 4) **Managerially sustainable** (will there be the local management capacity to carry on the work when the PCV leaves?).

Needs Assessment, falls within the Community Identification and Needs Assessment stage of the intervention logic. In the QCA analysis Needs Assessment is referred to as NEED, and NEED was calibrated to “1” representing the presence or to “0” representing the absence of a systematic formal or informal needs assessment being conducted before implementing the small project. The remaining nine conditions with their locations in the intervention logic, acronym used in the QCA, and their calibration are also presented.

The Logic Model displays the planned conditions during different stages of the project to achieve a successful outcome. However, as Cartwright et al. (2020) highlights, in addition to planned conditions, contextual support factors are often needed to help the functioning of these conditions; and identifying support factors is more likely to occur in retrospective evaluations since staff have the advantage of being able to reflect on the contextual factors that were critical for success. Of the ten conditions identified by staff and used in the QCA analysis, six conditions were specified in the Logic Model [CBCORE, NEED, TRAINER, START, LENGTH] and four conditions represented contextual support factors that supported the functioning of the planned conditions [ADM, SUPP, FOLLACT, INCENT].

Although QCA can be used with a much larger number of cases, it is essential to have in-depth knowledge of each case to reliably identify presence or absence of each condition in the project and, thereafter, accurately interpret the QCA results. A general recommendation (Ragin, 2008; Schneider and Wageman, 2012) is to include a mid-size number of cases which ranges from 10 and 50 cases. Thus, the evaluation team used purposive sampling approach in selecting 30 cases for QCA, ensuring that the sample included a broad range of implementation years, sectors, and performance outcomes.

Table 3 below presents the 30 SPA Program projects selected for inclusion in the QCA out of the total of 270 small projects. A total of 18 projects are from the category of Unsuccessful, that is, characterized by the absence of the desired outcome (OUTCOME = 0). Of these, 8 projects received the lowest rating in the Poor category and 10 projects had an average rating in the Adequate category based on the Performance Rubric. The 12 projects representing the presence of the outcome, Successful (=1), include 8 projects with the highest rating in the Excellent category and 4 projects with average ratings in the Good category.

To begin the process of identifying necessary and sufficient conditions of the outcome, a crisp-set data matrix was constructed, which is presented in Table 4. The first step in QCA is to identify necessary conditions, or in other words, those conditions that are consistently associated with the outcome (a successful small project). The data matrix rows list set membership scores for one case on each condition and the outcome, with ones indicating presence and zeros indicating absence of the condition.

2.5.1. Assessing necessary conditions

To identify necessary conditions, the evaluation team used fsQCA software (Ragin & Davey, 2016), which produced necessity-consistency scores that ranged from “0” to “1”, with “0” indicating no consistent association with outcome and “1” indicating complete association (necessity) with the outcome. Necessity-consistency is based on the number of cases that have a specific condition and a Successful outcome divided by the total number of cases that have a Successful outcome. In other words, the necessity-consistency score shows what percentage of successful cases had this condition present.

Table 5 below presents the full results of the Analysis of Necessity Conditions based on the crisp-set data matrix above. A consistency score of 1.0 means the condition yielded perfectly consistent scores with the outcome; thus, all 12 Successful SPA projects had this condition present in the expected direction of the program’s intervention logic; that is, their presence was associated with a successful outcome. Three of the ten conditions yielded perfectly consistent scores with a Successful outcome: 1) ADM (PCV received administrative and logistical support for their project), 2) SUPP (PCV received substantial support from an

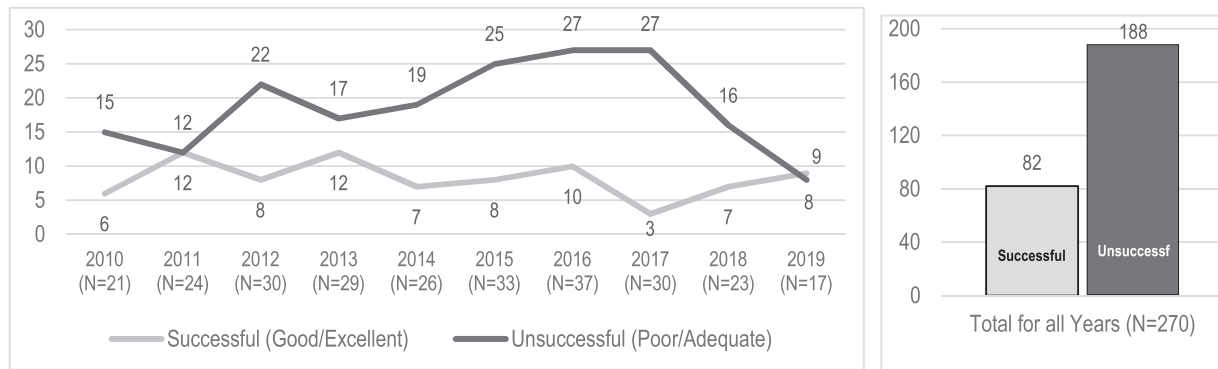


Fig. 2. Number of Successful and Unsuccessful Small Projects by Year and Totals.

Table 2
Location in Intervention Logic, Acronym, and Calibration of Ten Conditions for QCA.

Intervention Logic	Key Conditions (Acronym used in QCA analysis)	Present in Logic Model	Calibration
Relevance & Appropriateness	Community Identification Needs Assessment	1. [NEED] Needs assessment conducted.	Yes 1 = Formal or informal assessment conducted. 0 = None or little formal or informal assessment conducted.
		2. [YOUTH] Beneficiary age group.	No 1 = Youth (<25 yrs) are primary beneficiaries. 0 = Youth (<25 yrs) are NOT primary beneficiaries.
	Project Design	3. [LENGTH] Length of project.	Yes 1 = Less than 6 months. 0 = 6 months or more.
		4. [START] PCV in 1st or 2nd year of assignment.	No 1 = Project start date after PCV's 1st year of service. 0 = Project start date in PCV's 1st year of service.
		5. [CBCORE] Capacity-building core objective of project.	Yes 1 = Capacity-building need is core objective. 0 = Capacity-building need is not core objective.
Efficiency	Inputs	6. [TRAINER] Who conducted capacity-building training, PCV or someone else.	Yes 1 = Qualified local counterpart actively involved in capacity-building. 0 = PCV conducts capacity-building.
	Activities	7. [ADM] PCV received administrative and logistical support.	No 1 = PCV received substantial support in administrative and logistical tasks. 0 = PCV received little to no support in administrative and logistical tasks.
		8. [SUPP] PCV received support from an influential community member.	No 1 = PCV received support from an influential local community member. 0 = PCV received little to no support from an influential local community member.
		9. [INCENT] Project had strategies and activities to increase beneficiary interest and engagement.	No 1 = Outreach strategies and incentive activities to engage beneficiaries. 0 = None or little outreach or incentive activities to interest or engage beneficiaries.
Effects & Sustainability	Outputs	10. [FOLLACT] Post capacity-building hands-on activities.	No 1 = Follow-up activities after completion of capacity-building component. 0 = No follow-up activities after completion of capacity-building component.

influential counterpart or community member), and 3) CBCORE (capacity building was the core objective of the project).

In the second column of Table 5, necessity-coverage scores represent the proportion of all projects that had the condition present and simultaneously were Successful. That is, in the sample, the ADM condition was present in 21 small projects of which 12 were Successful, thus a coverage score of 0.57 (12 successful of the total of 21 projects with ADM), which indicates that administrative and logistical support was necessary (consistency =1) but not sufficient by itself (nine unsuccessful small projects had ADM support).

The condition, NEED (having conducted a needs assessment) is not perfectly associated with a successful project, with a consistency of 0.75, nonetheless, NEED represents, according to Befani (2016), "imperfect necessity... If the odds are calculated, we can say that a 75% necessity score means that the presence of the condition is needed three times more than its

absence." The consistency score of 0.75 indicates that NEED was present in 9 of the 12 Successful projects. Also, NEED had the highest Coverage score, indicating that 9 of the 11 projects (82%) that conducted a needs assessment were Successful and 16 of the 18 unsuccessful small projects (89%) had not conducted a needs assessment, which points to the empirical importance of NEED for producing an outcome (Schneider and Wageman, 2012).

In summary, the Analysis of Necessary Conditions did not show any single condition to be both necessary and sufficient in producing a successful small project. However, a conjunction or combination of two or more necessary conditions may be sufficient for the outcome, which was examined in a Sufficiency Analysis.

2.5.2. Assessing sufficient conditions

Sufficiency analysis attempts to answer the question, "what

Table 3
List of Projects (cases) Used in Qualitative Comparative Analysis.

Project Performance Rubric Category	# of SPA projects selected (total # of projects)	Project Selection Criteria	Year, Project Description, and ID # Selected for QCA
Poor	8 (of 31)	Lowest rating in poor category	2010 - Youth camp (ID #3) 2012 - Science lab (ID #6) 2013 - Sport/health (ID #9) 2014 - School library (ID #11) 2014 - English/Technology classroom in a public school (ID #13) 2015 - Community center (ID #15) 2015 - Economic development/Employability (ID #16) 2018 - English/Technology classroom in a public school (ID #27)
Adequate	10 (of 157)	Middle rating in Adequate category	2010 - English/Technology classroom in a public school (ID #1) 2013 - Economic development/Employability (ID #21) 2014 - Economic development/Employability (ID #10) 2014 - Youth camp (ID #14) 2014 - Economic development/Employability (ID #4) 2017 - Youth camp (ID #23) 2015 - Literacy development (ID #18) 2015 - Economic development/employability (ID #17) 2017 - English/Technology classroom in a public school (ID #24) 2017 - Youth center (ID #25)
Good	4 (of 61)	Middle rating in Good category	2010 - Community center (ID #2) 2011 - English/Technology classroom in a public school (ID #5) 2016 - English/Technology classroom in a public school (ID #22) 2016 - Youth empowerment (ID #29)
Excellent	8 (of 21)	Highest rating in Excellent category	2013 - Economic development (ID #8) 2014 - Economic development/Employability (ID #19) 2016 - English/Technology classroom in a public school (ID #30) 2017 - Youth camp (ID #26) 2016 - Economic development/Employability (ID #20) 2019 - Public health (ID #12) 2019 - NGO development

Table 3 (continued)

Project Performance Rubric Category	# of SPA projects selected (total # of projects)	Project Selection Criteria	Year, Project Description, and ID # Selected for QCA
			(ID #7) 2019 - Youth empowerment (ID #28)

combination of causal conditions are sufficient for the outcome?” (Schneider & Wageman, 2012) by examining conjunctural causation; that is, conditions do not necessarily exert their impact on the outcome in isolation from one another but sometimes must be combined to reveal causal patterns. Sufficiency analysis was used to identify the combination of conditions (or packages) sufficient for a successful small project among these cases. To examine the sufficient combinations or packages of conditions, a logical minimization process was applied to the crisp-set data. The goal of logical minimization is to produce a “truth table” that contains all possible combinations of conditions and then, using a Boolean minimization process, identify sufficient combinations of conditions. As mentioned earlier, with 10 causal conditions, there are 1024 possible causal combinations, the table below presents the combinations of conditions that were present in the cases included in the QCA analysis.⁴ As shown in Table 6, truth table rows describe the outcome for each possible combination of present and absent conditions, for all cases that have that combination.

The logical minimization process provides three sufficient solutions: simple, intermediate, and complex.⁵ Using fsQCA software, Table 7 presents the three sufficient solutions, or causal packages, that produce the outcome (successful projects) among the 30 small projects analyzed.

The simplest solution includes two conditions, NEED*TRAINER, which refers to projects that conducted a needs assessment AND (*) had a local counterpart conducting the capacity-building trainings rather than the PCV. This solution has a perfect sufficiency (consistency score = 1.0) meaning that all projects that conducted a needs assessment AND had a qualified local counterpart conduct capacity building training were Successful. The Raw Coverage score of 0.5 indicates that this causal package was present in 6 of the 12 Successful projects. The Unique Coverage score of 0.0 indicates that this solution, of two conditions, does not exclusively explain the outcome.

Two complex solutions were identified. The first complex solution includes either the presence or absence of the nine of the ten conditions, which was ~FOLLACT or the absence (indicated by the tilde symbol ~) of follow-up activities after the capacity-building trainings, ~YOUTH or the absence of a focus on youth in the capacity-building trainings, ~INCENT or the absence of a designated incentive mechanism integrated in the project design to motivate more engagement from the community, CBCORE or the presence of capacity building as a primary objective of the project, TRAINER or the presence of a local counterpart as a trainer for the capacity building activities, SUPP or the presence of substantial support from an influential community member, ADM or presence of administrative and logistical support, and ~START or the absence of project starting in PCV’s second year of service or in other words, the project started in PCV’s first year of service.

In contrast to the first complex solution, the second complex solution did not include capacity building follow-up activities [FOLLACT] but does include the project lasting less than 6-months [LENGTH]. Both complex solutions had perfect sufficiency (consistency score = 1.0)

⁴ Based on the formula, 2^k (Kent, 2008), or in this evaluation, 2¹⁰ equals 1024 combinations.

⁵ A simple solution is also referred to as a “parsimonious solution” and a complex solution is also referred to as a “conservative solution” (Schneider and Wageman, 2012).

Table 4
Crisp-Set Data Used in QCA.

Case ID	START	LENGTH	ADM	SUPP	NEED	TRAINER	CBCORE	INCENT	YOUTH	FOLLACT	OUTCOME
Project 1	1	1	1	1	0	0	0	0	0	0	0
Project 2	1	0	1	1	0	1	1	0	0	0	1
Project 3	1	1	0	0	0	0	1	0	1	0	0
Project 4	0	1	1	1	0	0	1	0	1	0	0
Project 5	1	0	1	1	0	0	0	0	0	0	1
Project 6	1	1	0	1	0	1	0	0	0	0	0
Project 7	0	1	1	1	1	0	1	1	0	1	1
Project 8	0	0	1	1	1	1	1	0	0	0	1
Project 9	1	1	0	1	0	0	1	0	1	1	0
Project 10	1	1	1	1	1	0	1	0	0	0	0
Project 11	1	1	0	0	0	1	0	0	0	0	0
Project 12	0	1	1	1	1	1	1	0	0	1	1
Project 13	1	0	0	0	0	0	0	0	0	0	0
Project 14	1	1	0	0	0	1	1	0	1	1	0
Project 15	1	0	1	1	0	1	0	0	0	1	0
Project 16	0	0	0	0	0	0	1	1	1	0	0
Project 17	0	1	1	1	0	1	1	0	0	1	0
Project 18	0	1	1	1	0	1	1	0	0	0	0
Project 19	0	0	1	1	0	1	1	0	1	1	1
Project 20	0	1	1	1	1	1	1	0	0	0	1
Project 21	0	1	1	1	0	1	1	0	0	1	0
Project 22	1	0	1	1	1	0	1	0	0	0	1
Project 23	1	0	1	1	0	1	1	1	1	1	0
Project 24	1	0	0	1	0	0	0	0	0	0	0
Project 25	1	0	1	1	1	0	1	0	1	1	0
Project 26	0	0	1	1	1	1	1	1	1	1	1
Project 27	1	0	0	0	0	0	0	0	0	0	0
Project 28	1	1	1	1	1	1	1	1	1	1	1
Project 29	1	1	1	1	1	1	0	0	1	1	1
Project 30	1	1	1	1	1	0	1	1	0	0	1

Table 5
Analysis of Necessary Conditions.

Outcome variable: OUTCOME (Good/Excellent performance rubric rating = 1)		
Conditions tested:		
	Consistency	Coverage
ADM	1.00	0.57
SUPP	1.00	0.50
CBCORE	0.83	0.52
NEED	0.75	0.82
TRAINER	0.67	0.50
START	0.50	0.32
LENGTH	0.50	0.35
FOLLACT	0.50	0.46
INCENT	0.33	0.67
YOUTH	0.33	0.36

meaning that all projects that had these conditions were also Successful. The Raw Coverage score of 0.17 indicates that these complex causal packages were present in 2 of the 12 Successful projects. The Unique Coverage score of 0.08 indicates that each of these solutions exclusively explain 8% of the outcome (1 of the 12 successful cases). The high consistency and low coverage scores indicate that although these complex solutions led to successful outcomes, they cannot explain most of the successful cases in the sample.

The intermediate solution contained the presence of five of the ten conditions and had perfect sufficiency (consistency score = 1.0) meaning that all projects that had these conditions were Successful. The Raw Coverage score of 0.42 indicates that this intermediate causal solution was present in 5 of the 12 Successful projects. The Unique Coverage score of 0.33 indicates that this solution exclusively explains 33% of the outcome (4 of the 12 successful cases).

Of the three solutions, the intermediate solution had the most robust model specifications. However, to determine which of these three causal packages to investigate further, the simple, intermediate, and complex solutions were presented to and discussed with the SPA Program staff. As recommended by Oana et al. (2021), the decision on selection of the

final solution should be based on not just model specifications but just as importantly on external knowledge, such as theory and experience, to see which causal package can be interpreted as conceptually meaningful. When the three causal packages were presented to the SPA Program staff, overwhelmingly they chose the intermediate solution, NEED*CBCORE*TRAINER*SUPP*ADM, as closer to their intervention logic and descriptive of their and PCV's experiences over the years than the simple solution or the complex solutions. SPA Program staff stated that the simple solution diminished the critical importance of community and PC administrative support during the life of the project for a project's success. As shown in the Truth Table, all 12 successful projects had SUPP and ADM present. Moreover, they stated that both complex solutions were 1) overly specific and represented too few cases and 2) would be too difficult for most PCVs to incorporate in their project planning.

As Ragin (2008) points out, intermediate solutions avoid both over- and under-simplification and therefore depict the most useful representation of reality. In addition, the intermediate solution has favorable properties compared to the simple (parsimonious) and complex (conservative) solutions (Schneider and Wageman, 2012).⁶ In Fig. 3 the intermediate causal package is placed within SPA Program's intervention logic framework. The first necessary condition in the causal package for a Successful project is conducting a thorough needs assessment among the community [NEED] to inform the project design, which ensured the project was relevant and appropriate to the community. However, although a needs assessment was necessary it was not sufficient alone for

⁶ Schneider and Wageman (2012, pg.175), "The intermediate solution has various important properties. It is a subset of the most parsimonious and a superset of the conservative solution. It is less parsimonious than the most parsimonious solution and more parsimonious than the conservative solution. Unlike the most parsimonious solution, it does not rest on difficult counterfactuals. Unlike the conservative solution, it incorporates theoretical hunches in the form of easy counterfactuals.... Intermediate solution terms therefore aim at striking a balance between complexity and parsimony, using theory as a guide as to which logical remainders should be assumed to have a link to the outcome."

Table 6
Truth Table for 10 Conditions Model.

ADM	SUPP	CBCORE	NEED	TRAINER	START	LENGTH	FOLLACT	INCENT	YOUTH	OUTCOME	Case ID
1	1	1	0	1	0	0	1	0	1	1	19
1	1	1	1	1	0	0	0	0	0	1	8
1	1	1	1	1	0	0	1	1	1	1	26
1	1	1	1	0	0	1	1	1	0	1	7
1	1	1	1	1	0	1	0	0	0	1	20
1	1	1	1	1	0	1	1	0	0	1	12
1	1	0	0	0	1	0	0	0	0	1	5
1	1	1	0	1	1	0	0	0	0	1	2
1	1	1	1	0	1	0	0	0	0	1	22
1	1	1	1	0	1	1	0	1	0	1	30
1	1	0	1	1	1	1	1	0	1	1	29
1	1	1	1	1	1	1	1	1	1	1	28
1	1	1	0	1	0	1	1	0	0	0	17, 21
0	0	0	0	0	1	0	0	0	0	0	13, 27
0	0	1	0	0	0	0	0	1	1	0	16
1	1	1	0	0	0	1	0	0	1	0	4
1	1	1	0	1	0	1	0	0	0	0	18
0	1	0	0	0	1	0	0	0	0	0	24
1	1	0	0	1	1	0	1	0	0	0	15
1	1	1	0	1	1	0	1	1	1	0	23
1	1	1	1	0	1	0	1	0	1	0	25
0	0	1	0	0	1	1	0	0	1	0	3
0	0	0	0	1	1	1	0	0	0	0	11
0	0	1	0	1	1	1	1	0	1	0	14
0	1	1	0	0	1	1	1	0	1	0	9
0	1	0	0	1	1	1	0	0	0	0	6
1	1	0	0	0	1	1	0	0	0	0	1
1	1	1	1	0	1	1	0	0	0	0	10

Table 7
Simple, Intermediate, and Complex Combinations Sufficient for Small Project Success.

Solutions	Sufficient Pathways*	Consistency	Raw Coverage	Unique Coverage	Cases
Simple	NEED*TRAINER Solution coverage: 1 Solution consistency: 1	1.0	0.50	0.00	#8, #12, #20, #26, #28, #29
Intermediate	NEED*CBORE*TRAINER*SUPP*ADM Solution coverage: 1 Solution consistency: 1	1.0	0.42	0.33	#8, #12, #20, #26, #28
Complex	~FOLLACT* ~YOUTH* ~INCENT*CBCORE*TRAINER*NEED*SUPP*ADM* ~START Solution coverage: 1 Solution consistency: 1 + ~YOUTH* ~INCENT*CBCORE*TRAINER*NEED*SUPP*ADM*LENGTH* ~START Solution coverage: 1 Solution consistency: 1	1.0	0.17	0.08	#20 #12

* A tilde symbol (~) refers to the “absence of.”

a successful small project because the needs assessment should focus on identifying capacities that need building or enhanced [CBCORE], which contrasts with small projects that focused solely on building or repairing infrastructure or supplying equipment to a classroom as the core objective. After the design of the capacity-building project, the next necessary conditions were receiving substantial support [SUPP] from one or more influential community members and delegating administrative and logistical [ADM] duties to a local counterpart during the input and activity stages. The last condition in the causal package, was having a local counterpart [TRAINER] conduct the capacity-building trainings rather than the PCV.

The placing of the causal package in the intervention logic framework appears to suggest a simple, sequential causal process; however, five conditions in the SPA Program causal package may not operate like a domino effect. Thus, as Befani (2012) asserts, “someone must actually put together the ingredients [conditions in the causal package] in a certain way to obtain the final effect.” Beach and Rohlfsing (2015) suggest four possible patterns of interaction among conditions within a causal package: 1) sequential occurrence, 2) simultaneous occurrence, 3) sequential and simultaneous, and 4) simultaneous and sequential occurrence. In Fig. 4, the five conditions in SPA Program’s causal

package are arranged in each of the four possible interaction patterns.

As shown in Fig. 4, the first pattern, sequential, suggests a sequence of each condition in the causal package leading to or triggering the next condition in succession. The second pattern, simultaneous, suggests all five conditions triggering and operating in a synchronized manner. The third pattern, sequential and simultaneous, suggests some conditions having a chronological interaction with the remaining conditions having a synchronized interaction. The fourth pattern, simultaneous and sequential, is the reverse of the previous pattern.

2.6. Causal process tracing

Although QCA identified a causal inference package, it is necessary to return to the cases that are members of this package to “unpack” the causal mechanisms behind this relational pattern (Illari, 2011; Russo & Williamson, 2007). As noted by Schneider and Rohlfsing (2013), “process tracing is an invaluable complement for QCA in order to discern the causal mechanisms behind a set relational pattern and further improve the theory and QCA model.” To unpack how the intermediate causal package is put together, two small projects were selected for causal process tracing.

The use of causal process tracing (CPT) in this evaluation is based on

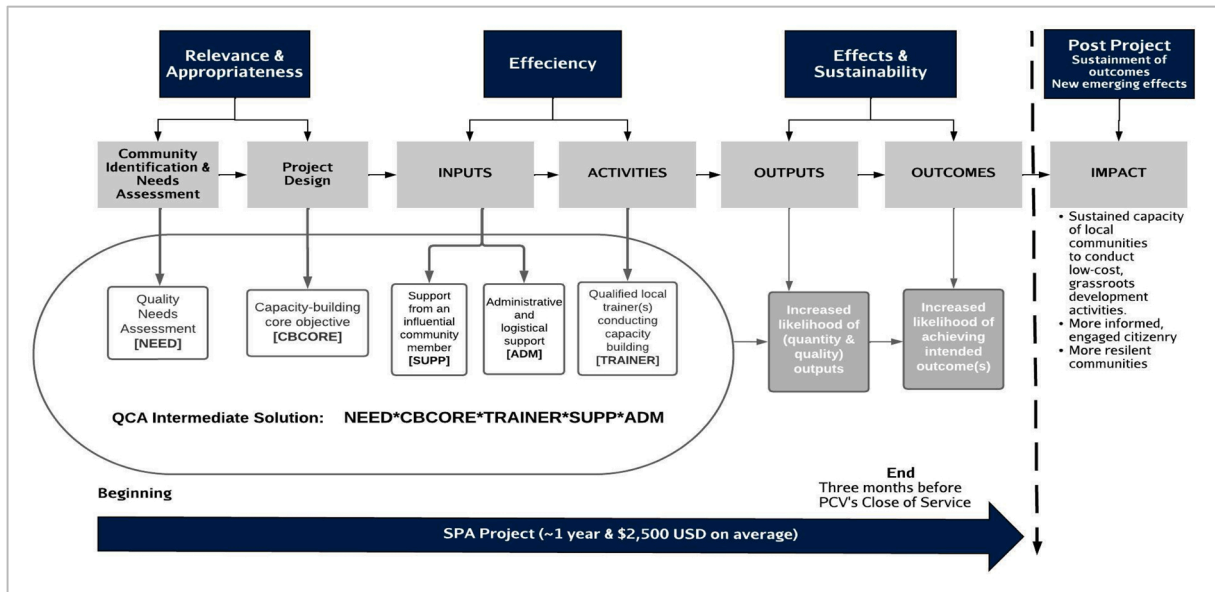


Fig. 3. Placing the Intermediate Causal Package Within the Intervention Logic Framework.

Sequential	NEED * CBCORE * SUPP * ADM * TRAINER	→	Successful Project
Simultaneous	NEED * CBCORE * SUPP * ADM * TRAINER	→	Successful Project
Sequential and Simultaneous	NEED * CBCORE * SUPP * ADM * TRAINER	→	Successful Project
Simultaneous and Sequential	NEED * CBCORE * SUPP * ADM * TRAINER	→	Successful Project

Fig. 4. Possible Patterns of Interaction Among Conditions in Intermediate Causal Package.

the results of QCA providing credible evidence to conduct within case analysis to unravel the mechanism through which it leads to the outcome in the case under study (Blatter & Haverland, 2014). Gerring (2007) refers to these as “pathway cases” that can provide insights into causal mechanisms. Following Machamer et al. (2000, pg. 3), these case studies investigated the empirically traceable patterns of interaction between entities and activities in the causal package. Entities that were examined included individuals (e.g., PCVs, PCV counterparts, influential community members, beneficiaries) and institutions (e.g., schools, local NGOs). Activities included events and actions such as conducting the needs assessment, establishing project objectives, planning and conducting trainings, and availability of administrative and logistical

support. In addition to investigating entities and activities, the context in which these entities and activities interacted was examined. That is, characteristics of the intervention and the Georgian context that possibly explain why the interaction of entities and activities helped achieve the outcome.

In addition, the QCA software, fsQCA, does not include a feature to identify individual cases causally comparable to a set of other positive (successful) cases (Falletti & Lynch, 2009). Thus, EvalC3 software (Davies, 2020), which has this feature, was used to select two small projects with successful outcomes as case studies to unpack the intermediate causal package. The View Cases menu provides a way to compare the “most similar” cases in the outcome group (successful

project), taking all attributes into account, not just those within the prediction model. Because the documentation for the successful Project #26 was relatively detailed, it was selected as the case to find a similar comparison. EvalC3 calculated Project #28 as the most similar project on all conditions. These small projects, to conduct causal process tracing, were a Youth Camp (#26) conducted in 2017 and Youth Empowerment (#28) conducted in 2019.

Case study 1: Youth Camp Project (ID #26) – The Youth Camp was a multi-regional camp held in western Georgia implemented by a well-known local organization working with youth for almost 20 years on issues of empowerment, leadership, and sexual and reproductive health and rights. In addition, in previous years this local organization had collaborated with PCVs to conduct youth camps.

[NEED] This local organization had a history of conducting high quality annual needs assessment research of young women throughout western Georgia that included surveys, focus group discussions, and in-depth interviews with youth, parents, and teachers. They analyzed these data by age groups, socio-economic status, and ethnicity. In 2017, one of the main findings of the annual needs assessment was increased need for information and counseling on health and reproductive rights among girls 13–17 years of age. Thus, the camp curriculum and activities developed by the organization was on general health and sexual reproductive health as well as understanding gender equality and civic engagement. In addition, the local organization had a strong outreach to and recruitment of disadvantaged young women and their parents.

[CBCORE] Given the conservative religious and cultural norms prevalent in Georgia, capacity-building of young women, especially in sexual and reproductive health was controversial with the potential for church and community opposition. Knowing local sensitivities to sexual and reproductive health among young women, the camp's curricula and activities were designed to build the capacity (awareness, knowledge, practice) of young women on general issues related to health, sexual and reproductive health, and gender.

[SUPP] For over two decades the local organization had organized and conducted numerous regional projects and had cultivated a positive reputation, was influential in regional youth issues, and had established connections with religious and political leaders in western Georgia. The local organization's longtime director fully supported the youth camp, which ensured readily available and accessible human, material resources, and logistical support when needed, thus greatly facilitating implementation of the camp.

[ADM] The local organization, due to their previous experience, was knowledgeable about different venue sites that could accommodate a youth camp, local food and equipment suppliers, medical services if needed, how to negotiate prices, and transportation and logistical issues. Also, the local organization's project coordinator staff worked closely with the PCV while the camp was ongoing to ensure any problems were quickly resolved.

[TRAINER] The local organization's trained peer educators conducted the training supported by a cadre of well-trained local young women as camp counselors and junior camp counselors. These peer educators, counselors, and junior counselors were all Georgian and, thus, used Georgian terminology for health and reproductive issues youth could easily understand, were familiar with the Georgian cultural norms and sensitivities, and anticipated challenges young women could face after the camp if they put into practice some of what they learned at camp about sexual and reproductive health.

Case study 2: Youth Empowerment Project (ID #28) – The goal of this project was to increase youth activism by providing them with different types of skills that they could use to advocate for addressing and solving relevant issues in the communities where they live. The project was implemented in a large city located in western Georgia.

[NEED] The PCV partnered with a well-known local youth-focused organization that had been active in the municipality for twelve years. The PCV and local organization, along with local youth, held a brainstorming session to discuss possible SPA projects; however, at the end of

this session, the main finding was that little was known about current challenges and problems of youth in the municipality to design a relevant project to meet a current need. Thus, the local organization, youth, and the PCV decided that the SPA project should be a comprehensive regional needs assessment of youth issues.

[CBCORE] Prior to conducting the comprehensive needs assessment, a curriculum of training were conducted for youth on 1) research methods and needs assessment techniques, 2) data analysis and report writing, 3) dissemination of needs assessment findings, and 4) advocacy techniques and civic activism on recommendations. After the training curriculum, the capacity-building continued in the form of hands-on activities, such as involvement in interviewing respondents, conducting data entry and analysis, report writing, and developing realistic recommendations. The needs assessment report was lauded at a conference attended by local NGOs, international NGOs, the local municipal government, and educational institutions. After this conference, the organization was invited to join the municipality's Youth Council.

[SUPP] Since the organization had been in operation for twelve years in the municipality, it had extensive contacts and good relations with a broad set of stakeholders in the municipality, such as local youth, parents, teachers, health officials, sports celebrities, law enforcement, and local NGOs.

[ADM] The PCV received substantial administrative and logistical assistance from the organization's Executive Director as well as designated administrative staff. The assistance included choosing vendors, negotiating prices, transporting materials, obtaining receipts, and providing on-going accounting of SPA program grant funds.

[TRAINER] Youth involved in conducting the needs assessment received several rounds of training and the first round of training was conducted solely by staff from the organization. The second round of training involved youth who completed the first round of training who then became peer educators assisting the primary trainers from the partnering organization.

In both case studies, the pattern of interaction among conditions in the intermediate causal package closely matched the sequential and simultaneous pattern. A sequential pattern occurred when after conducting a needs assessment [NEED] a relevant capacity building objective was established for the project [CBCORE] in the project design phase. After the project began, then a simultaneous pattern occurred with the synchronization of having the project supported by one or more influential community members [SUPP], the PCV receiving substantial administrative and logistical support [ADM], and a well-qualified and culturally sensitive trainer [TRAINER] conducting the capacity building trainings and workshops. Each of these three conditions were simultaneously supporting each other. Fig. 5 below represents what Cartwright et al. (2020) would call a local-setting process Theory of Change or in this evaluation local-level Logic Model.

2.7. Contextual Features Linking Conditions in the Intermediate Causal Package

Pawson and Tilley (1997) define contexts as the pre-existing institutional, organizational, and social conditions that can enable, slow down or derail actors, events, and activities. Falletti and Lynch (2009) refer to contextual conditions as “*relevant aspects of a setting (analytical, temporal, spatial, or institutional) in which a set of initial conditions leads ... to an outcome of a defined scope and meaning via a specified causal mechanism or set of causal mechanisms*” ...merely conditions that have to be present for a relationship to work in a particular manner.” The power of the causal package sufficient for a successful outcome, identified through the QCA, is influenced by the contextual characteristics present in rural Georgia which may or may not be present in other contexts where the SPA program is implemented. To illustrate the contextual conditions that shaped the causal package discussed in this paper, below each of the QCA conditions is discussed in the context of Georgia's socio-cultural background. Observations on the workings of each conditions were

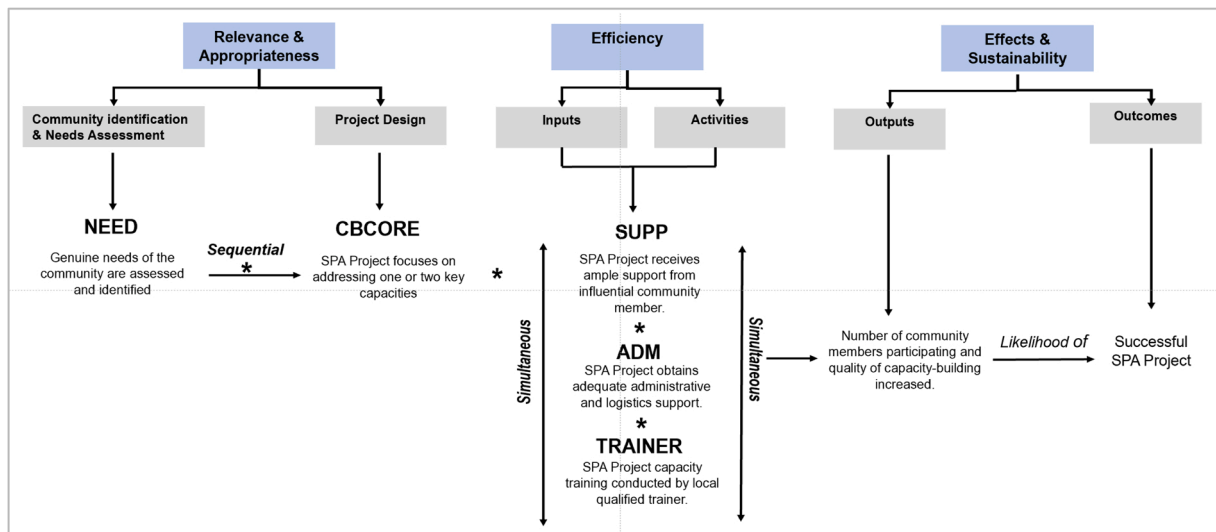


Fig. 5. Program’s Intervention Logic with Intermediate Sequential and Simultaneous Causal Package.

gleaned from the interviews with PCVs, PCV’s community counterparts, and beneficiaries.

NEED – PCVs often spoke about the politeness of local communities which made it difficult to understand their true interests and motivations. When PCVs communicated potential project ideas and wanted to gauge local support for that idea, many nodded in agreement and expressed support but PCVs struggled to tell how genuine that support was. Other times, PCV were approached by community members who said they knew what community youth needed, such as a principal of a local school who said that new sports equipment, even though it only benefitted one school, only youth with an interest in sports, and with little to no capacity-building. Most rural communities in Georgia have little to no experience of conducting needs assessments before designing a project; rather, projects are often initiated by one or more influential community members to serve the needs they deem to have the highest priority. Convincing a local community of the importance of a systematic needs assessment, particularly due to the short timeframe a PCV had to implement a small project, was often a challenge and in many small projects was not conducted, conducted poorly, or the “needs assessment” merely represented the desire of a local community member.

CBCORE – Financial hardships and a decrepit infrastructure experienced by many communities in Georgia often led community members to prefer small projects that involved obtaining equipment and supplies or infrastructure improvement. When a needs assessment was conducted in a quick and simplistic manner, such as a straightforward question, “what is needed,” all too often community members and local organizations prioritized immediate renovations (e.g., school gym) and purchase of equipment (e.g., multimedia equipment for a classroom). Since one of the main objectives of the SPA Program was capacity-building, this required the needs assessment to identify relevant and appropriate capacity-building needs that would provide longer-term, sustainable solutions for youth in the community.

ADM – In countries with a strong informal economy, such as Georgia, it can be difficult for a young PCV to navigate the local market and identify reputable suppliers, obtain the necessary accounting documentation, get access to resources at a fair price, keep suppliers accountable for the delivery of services, monitor the quality of products and services. Although PCVs adapt well to their local community, nonetheless, navigating an unfamiliar and, at times, a puzzling procurement system can be a difficult challenge. Thus, assistance from local counterparts in dealing with administrative and logistical aspects of a small project is critical, especially during start-up and implementation when delays can jeopardize a short-term project.

SUPP – In rural communities in Georgia, there is a strong deference to local informal and formal leaders, whether religious leaders or various officials (government, health, education). Having the support of one or more of these leaders and officials is often crucial to accessing scarce resources, securing the support of other local individuals to assist the project, gaining trust from the local community, promoting the project, and occasionally recruiting project participants. Some rural communities may show certain distrust towards outsiders, especially foreigners, which makes it difficult for PCVs to engage community members in their projects. Support from a locally influential person can be pivotal for gaining public acceptance and maintaining public engagement in the project.

TRAINER – Many capacity-building activities initiated by PCVs push cultural boundaries and motivate participants to change conventional norms and behaviors. For example, this can involve training and motivating teachers to incorporate more technology in their teaching, and motivating traditionally disengaged youth to become more interested and engaged in community issues, teaching young girls about their reproductive rights, and so forth. The success of such initiatives demands a sensitive socio-cultural approach, a good understanding of the local context, and clear communication, which can be difficult for PCVs in a foreign language. Having a well-qualified Georgian trainer that knows the language, understands socio-cultural barriers, and can motivate participants is essential.⁷

2.8. Explaining successful projects with several intermediate causal package conditions absent

Perfect subset relations are rare in case-level data (Ragin, 2008). Since the intermediate solution covers five of the twelve successful projects (one subset of a successful outcome), we examined the remaining seven successful projects, we call a “unique cases” subset, to understand how they achieved success with some conditions in the intermediate solution absent. Projects #2 and #19 dealt with youth employability for the rapidly developing tourism industry in Georgia. Neither of these two projects had conducted a needs assessment, nonetheless, youth employment is consistently seen as a prevailing need in public opinion polls and by the government in Georgia. The PCVs of these two projects cited these sources and decided not to expend

⁷ The United States State Department’s Foreign Service Institute places the Georgian language in Category IV (exceptionally difficult for English speakers) out of its five categories of language difficulty rankings.

resources (i.e., time and budget) to reconfirm a prevailing need (employment) and resolution (tourism guides). Although basing a project design on a general prevailing need is a regular recipe for success, in some cases, it can provide good guidance if the conditions of a qualified local trainer and capacity-building are present.

Projects #7, #22, and #30, did not use a local, qualified expert for capacity-building, which was conducted by the PCV. The interviews conducted during the evaluation showed that some PCVs struggled to engage participants when conducting capacity building in Georgian language, which is why a local qualified trainer increased the chances of a successful project. However, some PCVs were successful despite these challenges because of their ability to acquire sufficient Georgian language skills, and awareness of cultural sensitivities, during training and while living in the community. The intermediate solution provides a recipe for success that can be applied to all PCVs without preconditions about their Georgian language skills.

Project #5 entailed the procurement of various technological equipment, including a computer, projector, and speakers, for an English language classroom in the year 2016. Despite the absence of a formal needs assessment, and a minimal emphasis on capacity-building, the project was aligned with the Georgian government’s initiative to promote the integration of technology in the classroom. Notably, when the project started, the teachers at the school were already participating in government-sponsored trainings that focused on technology use in English language education. As such, the project supplemented an already recognized need within the school, with capacity-building primarily provided by the government, and not by local entities. The evaluation team’s site visit in 2021 revealed that the equipment in the technology classroom was still being actively utilized by teachers for English language classes.

Project #29 had all conditions of the intermediate solution present except a focus on capacity-building. Instead, addressing one of the top issues in the needs assessment, this project focused on challenging cultural norms that restricted the participation of girls in organized sports. Although no specific capacity-building training was included in the project, nonetheless, participating girls and families reported increased levels of confidence and self-esteem after the project. Typically, small projects aimed at changing deeply ingrained cultural practices have a low likelihood of success; however, this particular project was locally driven, owned, and sustained well beyond its completion. [Table 8](#).

2.9. Conditions associated with unsuccessful (failed) small projects

Another aspect of causal complexity examined in QCA is causal asymmetry, which is the possibility that the conditions leading to the presence of an outcome may be different than those leading to the outcome’s absence ([Schneider & Wageman, 2012](#)). Specifically, the conditions that predict the presence of success may not be the same conditions that predict the absence of success in small projects. Because of the possibility of asymmetrical causation in set-theoretical relations, a separate analysis of the absence of the outcome is needed.

A separate QCA analysis of the eighteen unsuccessful projects (OUTCOME=0), [Table 9](#) shows the presence and absence of all conditions and the conjunction of ~NEED* ~INCENT was sufficient to predict

Table 8
Explanation of the Unique Cases Subset of Successful Outcome (absence of several intermediate solution conditions).

Case#	ADM	SUPP	NEED	TRAINER	CBCORE	Explanation
Project 2 - Community center	1	1	0	1	1	Project based on recognized prevalent need instead of conducting a needs assessment
Project 19- Economic development/ employment	1	1	0	1	1	
Project 7 - NGO development	1	1	1	0	1	PCV gained sufficient Georgian language skills
Project 30 - English technology classroom	1	1	1	0	0	
Project 22- English technology classroom	1	1	1	0	1	
Project 5- English technology classroom	1	1	0	0	0	Project supported existing government teacher training
Project 29 - Youth empowerment	1	1	1	1	0	Project focused on changing cultural norms rather than specific skills training

Table 9
Consistency and Coverage for Presence and Absence of Conditions Associated with Small Project Failure.

Outcome variable: ~OUTCOME (Successful=0)		
Conditions tested	Consistency	Coverage
~NEED	0.89	0.84
~INCENT	0.89	0.67
START	0.72	0.68
SUPP	0.67	0.50
LENGTH	0.61	0.65
CBCORE	0.61	0.52
~YOUTH	0.61	0.58
~FOLLACT	0.61	0.65
~TRAINER	0.56	0.71
ADM	0.50	0.43
~ADM	0.50	1.00
TRAINER	0.44	0.50
~LENGTH	0.39	0.54
~CBCORE	0.39	0.78
YOUTH	0.39	0.64
FOLLACT	0.39	0.54
~SUPP	0.33	1.00
~START	0.28	0.45
NEED	0.11	0.18
INCENT	0.11	0.33
Conjunction: ~NEED* ~INCENT	0.82	0.78

14 of the 18 (or 78%) unsuccessful projects. That is, the absence of a needs assessment AND (*) the absence of activities and incentives to increase engagement from community members as a pathway to project failure.

With the absence of an adequate needs assessment, in which the community identifies their needs and project activities are tailored to address them, often incentives are required to generate interest and participation among the community in the project. However, without such incentives, community members may lack the motivation to contribute or participate in the project. There were projects when participation incentives were not included, but because the project was well-grounded in local needs the project succeeded. Conversely, there were cases where a needs assessment was lacking but useful incentives (food, publicity, recognition, fun activities) were effectively designed to keep engagement high and the project succeeded. In the absence of both conditions, project success (community ownership and sustainability) was more likely to be absent. Thus, two conditions could explain the absence of success in the majority (78%) of unsuccessful projects, whereas five conditions were needed to explain the presence of success in 42% of the successful projects.

3. Summary

Based on the small project performance rubric, collaboratively developed by the evaluation team and the SPA Program staff, 31% (or 82) of the 270 small projects implemented from 2010 to 2019 were successful. Of the ten conditions proposed to be influential in the success of a small project, QCA identified five conditions that were necessary but not sufficient alone. Sufficiency was achieved only when the five conditions were combined into a causal package. This causal package was

sufficient for a successful small project despite the age group of beneficiaries, the length of the small project, whether incentives to beneficiaries were provided, and whether follow-up activities occurred. CPT showed that when unpacked, the causal package involved the sequencing of a small project's initial activities of conducting a needs assessment followed by capacity-building and training plan based on the needs assessment. A simultaneous, mutually reinforcing relationship occurred during the implementation of the small project, which involved the PCV receiving support from an influential community member as well as administrative and logistical support from the SPA Program and the local community, as having the capacity-building training conducted by a qualified Georgian rather than by the PCV. In other words, the "productive continuity" (Machamer et al., 2000) of the program's intervention logic to achieve a successful small project came from a causal package of five conditions executed in proper sequential and simultaneous manner.

These findings were made possible because a set of theory-driven methods supported each other; specifically, the program's intervention logic was documented and graphically illustrated to guide the evaluation, a performance rubric was developed to determine what a successful small project looked like, cross case comparison, using QCA, identified a set of conditions that were necessary and sufficient for a successful outcome, and case studies, using CPT unpacked the causal package showing how it is put together and operates.

The causal package identified through this approach is highly context dependent. The causal package's effectiveness is based on the socio-cultural context in Georgia during these ten years or contextualized causality. The findings of this evaluation are meant to support current and future SPA Program effectiveness in Georgia.

3.1. Limitations of evaluation

All approaches and methods have limitations in answering evaluation questions. One of the main limitations for this retrospective evaluation was the challenges of examining possible causal packages over the ten years. Verweij and Vis (2021) state that the lack of a temporal dimension is one of the major drawbacks of QCA. Perhaps, the intermediate causal package evolved or changed over ten years. Although there are attempts to incorporate a time dimension in standard QCA, no approach has yet matched the level of complexity and sophistication that good case study research and its handling of time provides. Due to time and budget limitations of the evaluation, and commissioner request, case studies were not conducted to examine whether or not the extent causal packages changed over time in the SPA Program.

The second limitation encountered was the degree to which the intermediate causal package can be generalized to US Peace Corps' Small Project Programs established in other countries. The degree to which evaluation findings can be generalized is an important aspect of relevance. However, George and Bennett (2005) and Stern et al. (2012) highlight the "limited" and "contingent" generalization of QCA findings. Limited generalization referring to cases with the same target population, context, historical, and cultural factors; thus, limiting the generalization of the intermediate causal package to the SPA Program implemented in Georgia. Contingent generalization, however, would entail examining the degree to which US Peace Corps' SPA Programs in approximately fifty other countries use a similar intervention logic and the context is relatively similar to Georgia.

3.2. Lessons learned

First, prior to this evaluation, the SPA Program lacked a clear illustration of the intervention logic. The intervention logic was later developed with SPA Program staff and is shown earlier in Fig. 1. Basing further analysis on the intervention logic, this evaluation provided an empirically verified course of action to increase the success of small projects in the SPA Program in Georgia, which represents what

Carwright et al. (2020) refer to as a "thickening process", that is, replacing abstract theoretical components with validated local referents for a relevant design and model, as shown earlier in Fig. 5. Based on the finding of this evaluation, we believe that the project design and management trainings that PCVs receive should include discussions about successful and unsuccessful projects and the key conditions that led to these outcomes. Furthermore, the support that Peace Corps staff provides to volunteers throughout the project implementation process should be aligned with the key identified conditions that comprise the intermediate solution package for a successful project.

Second, all too often programs face budget limitations and thus program staff may be tempted to focus on one or two key components of the intervention logic in hopes that if these autonomous components are done well this will increase the likelihood of success. However, as shown in this evaluation, achieving success depended upon the conjunction of necessary conditions working together. Therefore, programs facing budget constraints should prioritize funding and implementing a sufficient package of necessary conditions rather than few individual components.

Third, we believe that the SPA Program staff in Georgia should incorporate the intermediate solution into its application and review process. As these conditions were shown to be key for project success, the application form may ask volunteers to present brief explanation of the plans for each of the five conditions (ex. what administrative hurdles do you anticipate in this project and what support do you have for dealing with them). This will help the SPA Program Review Committee to confirm that the PCV understands and has plans in place for all five conditions.

Author agreement

This article is an original work by the authors, has not been published before, and is not under consideration for publication elsewhere.

Permission note

Not applicable.

Funding

This evaluation was supported by the United States Agency for International Development (USAID) and US Peace Corps Office in Georgia.

CRedit authorship contribution statement

Larry Dershem: Conceptualization, Methodology, Analysis, Visualization, Writing. **Maya Komakhidze:** Conceptualization, Methodology, Data collection, Analysis, Writing. **Mariam Berianidze:** Data collection, Analysis, Review.

Acknowledgements

The authors would like to acknowledge the contribution of the US Peace Corps staff in Georgia, former PCVs in Georgia, and the extremely helpful PCV counterparts and community members who willingly gave their time to support this retrospective evaluation. In addition, the authors would like to thank a very helpful reviewer who made insightful comments on explaining successful projects in which the five conditions of the intermediate solution were not present as well as presenting findings regarding the unsuccessful projects. Finally, would like to acknowledge Margalita Japaridze who provided logistical support for this evaluation and to John Mitchell who provided substantive and practical recommendations on improving the manuscript.

References

- Beach, D., & Rohlfing, I. (2015). Integrating cross-case analyses and process tracing in set-theoretic research: Strategies and parameters of debate. *Sociological Methods & Research* (Vol.47, Issue 1), 3–36.
- Befani, B. (2012). Models of Causality and Causal Inference. *DFID Background paper*. Department for International Development.
- Befani, B. (2016). Pathways to change: Evaluating development interventions with qualitative comparative analysis (QCA). *Expertgruppen for Bistandsanalys*. Elanders Sverige AB, Stockholm.
- Blatter, J. and M. Haverland. (2014). Case Studies and (Causal-) Process Tracing. Chapter 4 in Engeli, I and C. Rothmayr. (2014) Comparative Policy Studies. Conceptual and Methodological Challenges. Houndsmills Basingstoke: Palgrave Macmillan.
- Campbell Collaboration. (2019). Campbell systematic reviews: policies and guidelines, Version 1.6, Series No. 1.
- Cartwright, N., Charlton, L., Juden, M., Munslow, T. and Williams, R.B. (2020). Making predictions of programme success more reliable. CEDIL Methods Working Paper. Oxford: Centre of Excellence for Development Impact and Learning (CEDIL).
- Chen, H. T. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage.
- Coryn, Chris, L., Lindsay, A. Noakes, Carl, D. Westine, & Daniela, C. Schroter (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199–226.
- Dalkin, Sonia Michelle, Joanne, Greenhalgh, Diana, Jones, Bill, Cunningham, & Monique, Lhussier (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, 10, 49, 2015.
- Davidson, E.J., Wehipeihana, N. and McKegg, K. (2011). The rubric revolution. Paper presenter at the Australasian Evaluation Society Conference. 1 September, Sydney, Australia.
- Davies, R. (2020). EvalC3 (2.4.42) [En; Excel]. (<https://evalc3.net/>).
- Donaldson, S. I. (2007). *Program theory-driven evaluation science*. New York, NY: Lawrence Erlbaum.
- Falletti, T. G., & Lynch, J. F. (2009). Context and causal mechanisms in political analysis. *Comparative Political Studies*, 42, 1143–1166.
- George, Alexander, & Bennett, Andrew (2005). *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Gerring, John (2007). Is There a (Viable) Crucial-Case Method? *Comparative Political Studies* (Volume 40, No. 3), 231–253.
- Grotzer, Tina A. and David N. Perkins. (2000) The Understandings of Consequence Project, Presented at the National Association of Research in Science Teaching (NARST). New Orleans, April 28- May 1,
- Hawkins, A. J. (2020). Program logic foundations: Putting the logic back into program logic. *Journal of MultiDisciplinary Evaluation* (Volume 16,(37), 38–57.
- Illari, P. Mc. Kay (2011). Mechanistic evidence: Disambiguating the russo-williamson thesis. *International Studies in the Philosophy of Science*, 25(2), 139–157.
- Kent, Ray. (2008). Using fsQCA: A brief guide and workshop for fuzzy-set qualitative comparative analysis. Department of Marketing, University of Sterling, Scotland. (<http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/workin-g-papers/2008/2008-10-teaching-paper-fsqca.pdf>) (accessed on 26 January 2023).
- King, Julian, Kate, Mc. Kegg, Judy, Oakden, & Nan, Wehipeihana (2013). Evaluative Rubrics: A Method for Surfacing Values and Improving the Credibility of Evaluation. *Journal of MultiDisciplinary Evaluation* (Volume 9,(Issue 21), 11–20.
- Machamer, P., Darden, L. and Craver, C.F. (2000). *Philosophy of Science*, Vol. 67(1):1–25.
- Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.
- Muco, Marta, Karen Aschaffenburg, Elizabeth Botkin, Giovanna Monteverde, Yvette Neisser, Thompson von Agner, Joseph Wantz, and Eugene, Wickett. (2018) Performance Evaluation of the Small Project Assistance (SPA) Program, FY13-FY17. Report to the United States Peace Corps, Washington, D.C.
- Oana, Ioana-Elena, Carsten, Q. Schneider, & Thomann, Eva (2021). *Qualitative comparative analysis using R: A beginner's guide*. Cambridge University Press.,
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, England: Sag.,
- Ragin, C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press.
- Ragin, C., & Davey, S. (2016). *Fuzzy-set/qualitative comparative analysis 3.0*. Irvine, California: Department of sociology. University of California.
- Rihoux, B. and Ragin, C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vols. 1–51). SAGE Publications, Inc.
- Rodgers, Patricia J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions Evaluation, 14; 29–48.
- Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2), 157–170.
- Schneider, C. Q., & Rohlfing, I. (2013). Combining QCA and process tracing in set-theoretic multi-method research. *Sociological Methods & Research*, 42(4), 559–597.
- Schneider, C. Q., & Wageman, C. (2012). *Set-theoretic methods for the social sciences A guide to qualitative comparative analysis*. Cambridge, UK: Cambridge University Press.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., and Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. Working Paper 38. Department for International Development. Glasgow, UK.
- United States Peace Corps. 2015. Peace Corps Georgia Small Grants Program: Guide for Peace Corps Georgia Volunteers and Their Counterparts. Tbilisi, Georgia.
- Verweij, Stefan and Barbara Vis. (2021). Three strategies to track configurations over time with Qualitative Comparative Analysis *European Political Science Review*, 13, 95–111.
- What Works Clearinghouse. (2008). Evidence Standards for Reviewing Studies. Accessed (https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_version1_standards.pdf).
- Larry Dershem** is Practice Lead in Monitoring and Evaluation with the Policy and Management Research Center in Tbilisi, Georgia. He has been involved with socio-behavioral research and program evaluation for almost three decades in west Africa, the Middle East, Caucasus, Central Asia, and Southeast Asia concerning HIV/AIDS, youth livelihoods, child nutrition, rural household adaptation to rapid socio-economic change, community development, and social and interorganizational networks. He holds a Ph.D. in Rural Sociology and an MS in Community Development from the University of Missouri-Columbia.
- Maya Komakhidze** is a Ph.D. student and a Graduate Researcher at the Lynch School of Education and Human Development at Boston College in Massachusetts, USA. Prior to starting the PhD program she worked as a researcher and an evaluator for five years in Georgia and internationally. She has worked as an internal as well as external evaluator and a researcher for international development organizations, research centers, NGOs, governmental institutions, and consulting companies. Her research and evaluation work focuses on education, measurement and assessment, quality assurance, community development, and gender equality.
- Mariam Berianidze** is a Researcher at PMC Research Center in Tbilisi, Georgia, where she has worked for more than three years specializing in qualitative data collection and analysis. She obtained a MS in Social and Political Science from Universitat Siegen, Germany, in 2018. Her latest research involves conducting needs assessments, case studies, and program evaluation among marginalize and vulnerable groups in Georgia.