AFRICA FOR RESULTS INITIATIVE

A PRIMER ON QUALITATIVE CAUSAL INFERENCE IN IMPACT EVALUATION DESIGNS: A FOCUS ON QUALITATIVE COMPARATIVE APPROACH AND THE PROCESS-TRACING METHODS

From the African Community of Practice on Managing for Development Results at the African Capacity Building Foundation (ACBF)

Case Study

N°50

SYNOPSIS

Randomized Control Trials (RCTs) provide the most credible approach for addressing causal attribution in quantitative evaluation designs. However, their recent proliferation should not overshadow qualitative approaches being developed for achieving the same. Africa is experiencing a wave of RCTs as evidenced by an increased number of studies and trainings accros the continent on the topic.

This AfCoP-AfriK4R case study was developed based on a review of published works. It builds on existing publications on the potential role of qualitative methods in the field of impact evaluation. The main objective of the paper is to reflect on the potential role of qualitative methods in analyzing causal claims when assessing development interventions.

Key findings: The paper notes the strengths of the randomized control designs while recognizing that most policy contexts and questions deal with a small samples that are not subject to the experimental methods used in RCTs. When evaluators are faced with small samples or single cases which require then to establish causal linkages between an intervention and outcomes, qualitative tools for causal inference can be useful for accomplishing this. The paper therefore discusses qualitative social science methods for assessing causal inference namely: the Qualitative Comparative Approach and the Processing-Tracing methods.

Key recommendations: The paper provided few examples that can serve as introductory material but there is need for detailed examples on how both Qualitative Comparative Approach (QCA) and Process-Tracing can be applied in qualitative impact evaluation designs. The recommendations are that: a) evaluators need to reflect carefully on the social phenomena under study before designing evaluation studies, b) QCA and Process-Tracing are complex qualitative tools where additional capacity building is commended to evaluators especially those based in Africa and resource limited states, and c) these skills gaps in applying the tools is an opportunity for capacity building organizations in Africa like ACBF and their partners to develop the capacities of the African Community of Practice members and other practioners on the continent.

Introduction

The paper seeks to provide a primer to qualitative causal inference within the field of impact evaluations by borrowing tools used in political and social sciences. It is motivated by the recent trend on the African continent where various capacity development efforts are focusing on the capacity of evaluators on the African continent to conduct Randomized Control Trials (RCTs). As the RCTs wave sweeps the continent, it is important to reflect on the qualitative methods that can be used in contexts where random assignment of participants to an intervention is not possible. The paper serves as a summary of the Qualitative Comparative Approach (QCA) and Process-Tracing and how they can be used in impact evaluations. Recent work

shows the increased focus on QCA and Process-Tracing in the evaluation of gender budgeting initiatives (Bamanyaki and Holvoet, 2016) and complex aid programs (Schmitt and Beach, 2015). Hence the need to focus on the application of these methods when conducting qualitative impact evaluations. These methods are complex. As such, this brief only serves as an introduction to this developing field and readers should use the suggested references to build their knowledge and skills in the use of both QCA and Process-Tracing.

Methodology

The paper is based on a review of published literature. It builds on existing literature on the potential role of qualitative methods in the field of impact evaluation. Three key texts were summarized in addition to references cited in the text. Ragin's textbook entitled: *The Comparative Method: Moving beyond qualitative and quantitative strategies* (Ragin, 1997) provides an extended discussion on the case oriented approaches for causal analysis. The key chapters include the second chapter that provides an extensive discussion on heterogeneity and causal complexity and chapter six, which discusses the basic concepts of the Boolean approach to Qualitative Comparison. Furthermore, George and Bennett's book: *Case Studies and Theory Development in the Social Sciences* (2005) provides guidelines for conducting case and theory oriented research. In particular, readers can focus on chapter 10 which provides guidelines on how to conduct Process-Tracing and Historical Explanation. The third key text is from Gary Goertz (2006) titled *Social Science Concepts: A Users Guide* which discusses the implications of theory in case selection and how the definition of concepts also tends to implicitly or explicitly define the causal mechanisms. These three texts are also key as they are followed by specific teaching cases, which allow the reader to see how the concepts explored in the text are applied in real world context.

What is causal inference?

Causal inference refers to the process of establishing the cause and effect relationship between an independent variable and a dependent variable (Jo, 2008). In addition, Holland (1986) describes three aspects of causal inference as involving: a) understanding the effects of causes, b) understanding the relative effects of causes and c) understanding that not everything can be a cause. In development interventions, for example, we are interested in understanding whether a given intervention has led to an observed outcome. The quantitative and qualitative approaches assume different cause-effect relationships which inform the design of research studies or evaluation designs.

RCT (Experimental Approach)

When done properly, an RCT is often referred to as the "golden standard" of causal attribution in assessing development interventions (Rubin, 2008). The key to conducting effective RCTs is randomization, which allows researchers to isolate the cause effect relationship by comparing the treated and control group on an outcome variable of interest. In RCTs, subjects are randomly assigned to two or more groups, usually referred to as an intervention group or a control group (Boruch, 1994). An intervention group is the one that receives a program intervention while the control group does not receive any intervention. The procedure involves collecting baseline data characteristics of interests for both the intervention and control group. The baseline is then followed by an intervention (or treatment) (Bloom, 2006). The next stages involve measuring the characteristics of interest again and compare the scores between the intervention and control groups. Since the intervention would have been applied to one group and not the other, the difference between the intervention and the control group is attributed to the intervention (treatment). For this conclusion to apply, people should have been assigned to either the intervention or control group in a random way. In other cases, randomization may not be possible and causality can be estimated using non-

experimental approaches (Barrett and Carter, 2010).

Non-experimental approaches

In some cases, programs may want to assess causal impact after the program has already started. However, in such cases it is difficult to randomly assign project beneficiaries to either the intervention or control group. In order to create another group to assess what would have happened to the group that does not get the treatment, referred to as the counterfactual, researchers can either use the regression discontinuity methods or propensity score matching. Regression discontinuity refers to the process of including individuals that are outside the eligibility cut off (Imbens and Lemieux, 2008). For example, if the program targeted groups with income below USD 1,000, the counterfactual group can be comprised of people earning USD 1,001-1,005. Propensity score matching on the other hand seeks to identify groups that have similar observable characteristics in order to create a counterfactual (Rosenbaum and Rubin, 1983). In both cases the treatment score are compared to the counterfactual to assess the impact of the intervention. In the last case, if a measure was collected before the intervention, this baseline data can also be compared to the scores obtained from the intervention groups.

In both experimental and non-experimental designs, the ability to assess causality is based on the difference between the intervention group and the control group. In the field of experimental designs, RCTs are considered to be the most accurate and reliable for assessing causality (Acland, 1979 and Rubin, 2008) but there are several challenges associated with the use of experimental designs in assessing causal attribution. First, the idea that researchers can hold variables and study them independent of external influence as in the case of RCTs is often disputed in development interventions. Second, RCTs may not be feasible due to the huge costs associated with implementing them (Jager et al., 2007). Third, RCTs may be unethical or inappropriate in certain contexts particularly when applied to development economics (Barrett and Carter, 2010). Fourth, in real world contexts, you may not have sufficient units to assign randomly to the intervention and the control group - a total sample size that is large enough to provide reliable statistics (Pogue and Yusuf, 1998).

One of the most cited weakness of the experimental design is the idea that one cause leads to an outcome when other intervening factors are held constant (Barrett and Carter, 2010). In this brief we seek to extend the idea of multiple causal relationships and discuss some of the alternative methods for testing cause-effect relationship in non-experimental ways. We focus on two qualitative methods of QCA (Ragin, 1997) and Process-Tracing (Collier, 2011) and their applicability in testing how development interventions affect development outcomes.

Alternative views of causation

The main ideas around alternative views of cause-effect relationships are threefold: a) there is no clearly demarcated cause of social phenomena, b) cause-effect relationships rarely occur in isolation, and c) the effects of causes tend to vary by context (Ragin, 1997). We first discuss how QCA and Process-Tracing view causal relationships and then proceed to describe the methods for testing evidence in each of the approaches.

Qualitative comparative analysis (QCA)

Charles Ragin's seminal work on QCA describes how social science questions tend to follow the statistical rules in framing their questions and how this often limits social researchers from answering pertinent social, political, and geographical questions about human nature (Ragin, 1997). Since questions are framed to follow methods and not vice versa, researchers then end up understanding the average effects as all samples have to be large enough to allow for statistical analysis. Ragin views social causations as multiple and conjectural –

meaning social phenomena is caused by more than one event conjectural because multiple conditions often combine to lead to an outcome. Therefore, if researchers and/or evaluators assume that cause-effect relationships are conjunctural, then the focus of the inquiry is about understanding how different phenomena combine to produce social outcomes unlike the assumed linear cause-effect relationships in experimental designs (Berg-Schlosser et al., 2009).

Process-Tracing

Process-Tracing is defined as "an analytic tool for drawing descriptive and causal inferences from diagnostic pieces of evidence often understood as temporal sequence of events or phenomena" (Collier, 2011, p. 824). Process-Tracing attempts to uncover the links between causes and observed social phenomena by examining multiple sources of evidence. These may include historical evidence, notes and transcripts in order to assess whether the causal mechanisms that are proposed in a theory or hypothesis do exist within a given case or observation (George and Bennett, 2005).

Robust tests of evidence

Both the QCA and Process-Tracing provide methods for testing causal claims and eliminating rival explanations. Each of the tests is briefly discussed in this section.

QCA relies on the Boolean logic for assessing multiple causations and Ragin (1997) proposes an extended version of Boolean Logic. Both methods are complex and this section will not do justice to the details provided in the main reference text. Readers are encouraged to review Chapter 6 and Chapter 7 of *The Comparative Method: Moving beyond qualitative and quantitative strategies* (Ragin, 1997). The basic idea with the Boolean Logic is comparing the cases logically across multiple variables. In doing so, one is able to identify patterns of multiple conjunctural causation. Ragin proposes a six-step approach for conducting basic Boolean inference which involves 1) using binary data (coding variables for absence (0) or presence (1), 2) present the data as a Truth Table, 3) Apply Boolean addition, 4) Apply Boolean multiplication, 5) Apply the combinatorial logic, and 6) Apply Boolean minimization (see also Thiem and Duşa, 2013). These logics allow one to assess whether causes are sufficient or necessary. Table 1 is a Truth-table that best illustrates the representation of variables and outcomes using QCA. The variables can take either a 0 or 1 and they are associated with either the presence or absence of an outcome. Table 1 provides an attempt to establish causal mechanisms for why elite capture occurs at local level by coding five mechanisms for presence or absence across four countries.

Table 1: An example of truth table used in QCA

Country	Mechanisms				
	Entitlements	Demands	Allocation Decisions	Elite preferences	Misappropriation
Botswana	0	0	1	1	1
Namibia	1	1	1	1	1
Zambia	1	1	1	1	1
Zimbabwe	0	1	1	1	1

Source: Adapted from Muyengwa, S., Child, B., & Lubilo, R. (2014). Elite capture: A Comparative Case Study of Meso-Level Governance in Four Southern Africa Countries. Eds. G. Barnes & B.Child. *Adaptive Cross-scalar Governance of Natural Resources. Routledge.*

The extended Boolean methods are meant to deal with issues of diversity and contradictions. Social phenomena inherently have a limited number of outcomes, for example some variables can only take two values such as gender. The second challenge of contradictions has to do with the fact that a same combination of causes may not lead to the same outcome (presence or absence of some social phenomena). In such cases, there is need to

further refine the cases by defining and looking for additional causes. These complexities have led to the development of software to help in QCA analysis.

The five main softwares provided on the COMPASS website[1] are: fs/QCA[2], Kirq[3], cna[4], QCA[5], and QCAGUI[6]. The full list of softwares can be found on the COMPASS website and additional material for conducting QCA.

Process-Tracing offers four ways of validating the strength of causal propositions namely: (1) straw-in-the wind, (2) Hoop Test, (3) Smoking-Gun, and (4) Doubly Decisive (Mahoney, 2000). Some techniques are stronger than others. However, generally each of the tests is used to a) establish that an event occurred, b) subsequent or antecedent chain of events then followed the initial event and c) that the first event was the cause of the latter. In addition, each of the tests is assessed on whether or not it is necessary for affirming causal inference or sufficient for affirming causal inference (Collier, 2011). A summary regarding the four ways of validation are presented in Table 2. Further details then follow thereafter.

Table 2: Process-Tracing tests

	No	Yes
Necessary for affirming causal inference (No)	***(1) Straw-in-the-wind-test***	***(3) Smoking-Gun***
	Passing**:** Affirms the relevance of hypothesis but does not confirm it Failing: Hypothesis is not eliminated, but slightly weakened **Implications for rival hypotheses:** Passing: slightly weakens them Failing: Slightly strengthens them	Passing: Confirms hypotheses. Failing: Hypotheses is not eliminated, but is somewhat weakened. **Implications for rival hypotheses:** Passing: substantially weakens them. Failing: somewhat strengthens them.
Necessary for affirming causal inference (Yes)	***(2) Hoop***	***(4) Doubly Decisive***
	Passing: Affirms relevance of hypotheses but does not confirm it. Failing: Eliminates hypotheses. **Implications for rival hypotheses:** Passing: somewhat weakens them. Failing**:** somewhat strengthens them.	Passing: Confirms hypotheses and eliminates others. Failing: Eliminates hypothesis. **Implications for rival hypothesis:** Passing eliminates them. Failing substantially strengthens.

Source: Adapted from (Collier, 2011, p. 825)

a. Straw-in-the-wind

The straw in the wind tests serve to weaken the alternative hypothesis but they do not provide sufficient evidence to either reject or accept a hypothesis and are considered the weakest of the four tests used in Process-Tracing. The tests often

[1] www.compass.org/software.htm
[2] Ragin, Charles, and Sean Davey. 2014. *fs/QCA [Computer Programme]*, Version 2.5. Irvine, CA: University of California
[3] Reichert, Christopher and Claude Rubinson. 2014. *Kirq [Computer Programme]*, Version 2.1.12. Houston, TX: University of Houston-Downtown.
[4] Baumgartner, Michael, and Alrik Thiem. 2015. "Identifying Complex Causal Dependencies in Configurational Data with Coincidence Analysis." *The R Journal* 7 (1):176-84.
[5] Thiem, Alrik, and Adrian Duşa. 2013. "QCA: A Package for Qualitative Comparative Analysis." *The R Journal* 5 (1):87-97.
[6] Dusa, Adrian. 2007. "User Manual for the QCA (GUI) Package in R." Journal of Business Research 60 (5):576-86.

involve using the available evidence to assess whether an event has occurred. For example, in development intervention, this may include evidence that a recipient received cash voucher. With the evidence that the individual benefited from program X, the evaluator can be confident in the program theory of change or hypothesized outcomes. However, even in the absence of a cash voucher receipt in our example, only raised doubt about the truth of a hypothesis but is not evidence to rule it out (Centre for Development Impact, 2015)

b. Hoop

Hoop tests provide necessary but not sufficient explanations for accepting a hypothesis. They are characterized as hoops because the hypothesis must "jump through the hoop" in order for it to be an acceptable explanation. To adapt the examples provided by Evera (1997), we may be asking in our example presented above, if the person stayed in the village at the time when the cash vouchers were distributed. It is necessary therefore to establish that the person stayed in the village at the time when cash vouchers were distributed but this is not sufficient to explain that cash vouchers led to a hypothesized outcome such as increased school attendance.

c. Smoking-Gun

As indicated in Table 2, passing the Smoking-Gun test confirms the hypothesis but it does not eliminate the alternative hypothesis but only serves to weaken it (Collier, 2011). The presence of a Smoking-Gun can provide evidence or suspicion that the events led to an outcome but without prior knowledge, one may not be able to link the evidence to an outcome. Collier (2011) also cautions against the selective use of these tests to confirm a hypothesis or eliminate rival hypotheses.

d. Doubly Decisive

The Doubly-Decisive test is considered the strongest of the four tests. This is so because passing the test confirms the hypothesis and eliminates other alternative explanations. For example, evidence of a photograph can be used to confirm that an event took place (and with additional features such as geo-location). As Ever notes, in most cases, such evidence is often lacking and the Doubly Decisive test can be achieved by combining the Hoop and the Smoking-Gun.

Recommendations and conclusions

The purpose of this AfCoP-AfriK4R case study has been to highlight alternative views to analyzing cause-effect relationships beyond the traditional intervention versus control design. The non-experimental designs for inferring cause-effect relationships adopt assumptions that are radically different from the experimental approach, and assume complex interactions between cause-effect relationships defined as multiple conjunctural causation. The paper provided few examples that can serve as introductory material but there is need for detailed examples on how both QCA and Process-Tracing can be applied in qualitative impact evaluation designs. The recommendations are that: a) evaluators need to reflect carefully on the social phenomena under study before designing evaluation studies, b) QCA and Process-Tracing are complex qualitative tools where additional capacity building is commended to evaluators especially those working in Africa, and c) the skills gap in applying these tools is a potential opportunity for capacity building organizations in Africa like ACBF and their partners to develop the capacities of the AfCoP-MfDR members and other practioners on the continent.

References

Acland, H., 1979. Are Randomized Experiments the Cadillacs of Design? Policy Anal. 5.

Bamanyaki, P.A., Holvoet, N., 2016. Integrating theory-based evaluation and process tracing in the evaluation of civil society gender budget initiatives. Evaluation 22, 72–90. doi:10.1177/1356389015623657

Barrett, C.B., Carter, M.R., 2010. The Power and Pitfalls of Experiments in Development Economics: Some Non-random Reflections. Appl. Econ. Perspect. Policy 32, 515–548. doi:10.1093/aepp/ppq023

Berg-Schlosser, D., De Meur, G., Rihoux, B., Ragin, C.C., 2009. Qualitative comparative analysis (QCA) as an approach. Config. Comp. Methods 1–18.

Bloom, H.S., 2006. The core analytics of randomized experiments for social research. N. Y. Manpow. Demonstr. Res. Corp.

Boruch, R.F., 1994. The Future of Controlled Randomized Experiments: A Briefing. Eval. Pract. 15, 265–74.

Centre for Development Impact, 2015. Straws-in-the-wind, Hoops and Smoking Guns: What can Process Tracing Offer to Impact Evaluation? Innnovation and learning in impact evaluation 10, 8.

Collier, D., 2011. Understanding process tracing. PS Polit. Sci. Polit. 44, 823.

Evera, S.V., 1997. Guide to methods for students of political science. Ed Cornell NY.

Goertz, G., 2006. Social science concepts: A user's guide. Princeton University Press.

Holland, P.W., 1986. Statistics and causal inference. J. Am. Stat. Assoc. 81, 945–960.

Imbens, G.W., Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. J. Econom. 142, 615–635.

Jager, K.J., Stel, V.S., Wanner, C., Zoccali, C., Dekker, F.W., 2007. The valuable contribution of observational studies to nephrology. Kidney Int. 72, 671–675. doi:10.1038/sj.ki.5002397

Jo, B., 2008. Causal inference in randomized experiments with mediational processes. Psychol. Methods 13, 314.

Mahoney, J., 2000. Strategies of causal inference in small-N analysis. Sociol. Methods Res. 28, 387–424.

Pogue, J., Yusuf, S., 1998. Overcoming the limitations of current meta-analysis of randomised controlled trials. The Lancet 351, 47–52. doi:10.1016/S0140-6736(97)08461-4

Ragin, C.C., 1997. The comparative method: Moving beyond qualitative and quantitative strategies. Univ of California Press.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Rubin, D.B., 2008. Comment: The design and analysis of gold standard randomized experiments. J. Am. Stat. Assoc. 103, 1350–1353.

Schmitt, J., Beach, D., 2015. The contribution of process tracing to theory-based evaluations of complex aid instruments. Evaluation 21, 429–447. doi:10.1177/1356389015607739

Thiem, A., Duşa, A., 2013. Boolean minimization in social science research: A review of current software for Qualitative Comparative Analysis (QCA). Soc. Sci. Comput. Rev. 0894439313478999.

Acknowledgement

This knowledge series is intended to summarize good practices and key policy findings on managing for development results. The views expressed in the notes are those of the author. AfCoP Knowledge products are widely disseminated and are available on the website of the Africa for Results initiative (AfriK4R), at: www.afrik4r.org/page/resources.

This AfCoP-MfDR knowledge product is a joint work by the African Capacity Building Foundation (ACBF) and the African Development Bank (AfDB). This is part of the knowledge products produced by ACBF under the leadership of its Executive Secretary, Professor Emmanuel Nnadozie.

The product was prepared by a team led by the ACBF's Knowledge, Monitoring, and Evaluation Department (KME), under the overall supervision of its Director, Dr. Thomas Munthali. Within the KME Department, Ms. Aimtonga Makawia coordinated and managed production of the knowledge product while Dr Barassou Diawara, Mr. Kwabena Boakye, Mr. Frejus Thoto and other colleagues provided support with initial reviews of the manuscripts. Special thanks to colleagues from other departments of the Foundation who also supported and contributed to the production of this paper. ACBF is grateful to the Africa Development Bank which supported production of this MfDR case study under grant number 2100150023544.

ACBF is also immensely grateful to Mr. Shylock Muyengwa, as the main contributor, for sharing the research work which contributed to the development of this publication. We also thank Prof G. Nhamo, Dr. Lyimo and Dr. A. Kirenga whose insightful external reviews enriched this knowledge product. The Foundation also wishes to express its appreciation to AfCoP members, ACBF partner institutions, and all individuals who provided inputs critical to completing this product. The views and opinions expressed in this publication do not necessarily reflect the official position of the ACBF, its Board of Governors, its Executive Board, or that of the AfDB management or board.