Review Paper

Qualitative comparative analysis for WASH research and practice

Jessica Kaminsky and Elizabeth Jordan

ABSTRACT

Qualitative comparative analysis is an established research method that has been underutilized in water, sanitation, and hygiene (WASH) research. It has immense potential for addressing the complexity inherent to WASH projects, and can produce robust and transparent results from intermediate or large numbers of cases. The method enables researchers and practitioners to blend quantitative and qualitative metrics to build more nuanced contextual knowledge, and is able to detect combinations of causal conditions that lead to outcomes of interest. This means that the method is uniquely positioned for building empirically founded theories of change that reflect contextual complexity. In this review paper we use hypothetical data and a review of the existing literature to showcase where and how the method can be productively applied in WASH research and practice.

Key words | hygiene, methods, QCA, sanitation, WASH, water

Jessica Kaminsky (corresponding author)
Department of Civil and Environmental Engineering,
University of Washington,
121H More Hall,
Seattle,
WA 98195,
USA
E-mail: *jkaminsk@uw.edu*

Elizabeth Jordan
USAID,
1300 Pennsylvania Ave, NW,
Washington,
DC 20523,
USA

INTRODUCTION

Recent years have seen the water, sanitation, and hygiene (WASH) community increase its focus on evidence-based approaches, monitoring, and evaluation. This move is intended to improve accountability and results for both donors and the communities where projects take place. In part, this trend towards measurement is a reaction to the most fundamental and important questions for global development: why is it that some projects succeed while others fail? And what, exactly, do we mean by success and failure in WASH projects? While occasionally there may be simple answers to these questions, more often the answers themselves are complex systems of dynamic, contextual factors and decoupled impacts (Meyer & Rowan 1977).

The gold standard for research methodology is often considered to be the randomized controlled trial (RCT), which emulates clinical trial research methods. For example, a handful of recent RCTs have tested the impact of community-led total sanitation (CLTS) methods (Clasen *et al.* 2014; Patil *et al.* 2014; Guiteras *et al.* 2015; Pickering *et al.* 2015). Like any tool, however, RCTs are not perfect. One practical problem is the considerable expense of implementation. Other methodological issues are common to any quantitative approach; closed-ended questionnaires allow statistical analysis but force individual responses into pre-determined schema that may or may not be appropriate. In a related methodological issue, quantitative methods prove relationships but struggle to discover how or why variables contribute to the outcome of interest. For example, why did the previously referenced and well-designed CLTS studies show differing impacts on outcomes like stunting, incidence of diarrhea, and rates of change in

doi: 10.2166/washdev.2017.240

latrine ownership? Different and complementary research methods – like the qualitative comparative analysis (QCA) method described in this paper – are needed to answer these important questions.

In contrast to statistical methods, traditional qualitative research approaches allow local knowledge to emerge from the data and are well suited to discovering how and why WASH interventions work in a particular context. However, qualitative findings cannot be statistically generalized, and the very nuance of the answers that qualitative methods generate can mean they are relatively difficult to communicate and apply in different contexts. Because of these differing strengths and weaknesses, high quality quantitative and qualitative research deeply complement each other. Each type of approach – and there are many methods within these broad categories – can answer different kinds of research questions. For example, the very complexity of factors discovered through qualitative cases may provide an explanation for why it is so difficult to statistically link WASH interventions and health outcomes (Schmidt 2014). Or, quantitative studies may statistically validate qualitative findings, discovering the importance of each factor relative to the outcome of interest.

QCA (Ragin 2008) is a research method that blends the strengths of qualitative and quantitative methods. QCA is a set-theoretic method that seeks combinations of causal conditions, or pathways, that lead to an outcome of interest. To do so, deep qualitative and quantitative case knowledge is explicitly represented by calibrated, quantitative measurements in a truth table. This truth table is then simplified using either Boolean algebra or fuzzy logic in a fully reproducible, generalized set theoretic analysis. The method lives between qualitative and quantitative analysis, and can handle either intermediate or large numbers of cases. While a relatively new research method (Ragin 2008) which was originally used in the areas of comparative politics and historical sociology, over recent decades it has made significant inroads to a wide range of research communities, including economics, management and engineering. QCA allows us to rigorously analyze different types, quantities, and combinations of qualitative and quantitative data; thus we suggest it is an important addition to the WASH toolkit.

QCA is founded in set theory. The simplest set relation is a subset – for example, water projects are a subset of WASH projects. This type of set relation defines both water projects (as one kind of a WASH project) and also defines WASH projects (as including, among other things, water projects). While definitional sets can be trivial, more interesting set relations emerge when researchers seek causal relationships between various phenomena; the causality claim, of course, is founded on theory and sector knowledge. For example, and as discussed below, there are both sustainable and unsustainable school sanitation projects, and we suspect there are reasons why projects turn out to belong to one or the other of these subsets of school sanitation projects. Causal conditions are the reasons that the researcher believes may influence the outcome of interest. While these are similar to independent variables in a statistical analysis, they do not take on many of the assumptions of variables in a statistical analysis. As we discuss below, set relationships are importantly different than correlational relationships, in part because of the underlying assumptions about the symmetry of theorized relationships between causes and outcomes (see Table 4 and the related discussion for an example of this difference). The thought structure of this paragraph parallels that of the first chapters of Ragin's (2008) book. The reader is referred to that book for more details on the fundamentals of set theoretic thought, which are not limited to the introductory examples provided here. Given the difference and utility of the QCA research method, this paper is a methodological contribution intended to describe how QCA may be useful to the WASH community.

QCA FOR WASH

In recent years a handful of researchers have begun to apply QCA to WASH research, which we define as research interested in drinking water supply, sanitation and hygiene for developing nations, communities, and households worldwide. While limited in number, the existing studies showcase the wide variety of applications in which QCA can be valuable. To identify the examples referenced here (which necessarily represent a subset of those examples present in the literature), we searched the literature for 'QCA' in

combination with either 'WASH,' 'water,' 'sanitation,' or 'hygiene.' We included papers that considered WASH projects as some cases in a larger dataset. These searches were carried out on article databases including Academic Search Complete, Engineering Village, Web of Science, the Environmental Science Collection, SCOPUS, and WorldCat. We also reviewed the first 100 results for each search on the GoogleScholar search engine. A limitation of this approach is that most non-academic publications are not archived in these databases. As such, for each combination of search terms we also reviewed the first 100 results returned from standard Google searches, as well as searching the SuSanA knowledgebase (http://www.susana.org/en/resources/library) for 'QCA'. This approach identified resources such as QCA-based evaluation protocols and toolkits (Annamalai *et al.* 2016; OpenIDEO 2016). This search identified 17 key documents, including four practitioner-published reports, two dissertations, and 11 academic journal articles.

The majority of the key articles we identified dealt with water, while a few treated sanitation and hygiene. The cases analyzed in the articles varied widely in scale. For example, several papers analyzed individual households (Spencer 2008; Kaminsky & Javernick-Will 2014) or schools (Chatterley *et al.* 2013, 2014), others analyzed development projects (Boudet *et al.* 2011; Santosh Kumar Delhi *et al.* 2012), and another analyzed public private partnerships for urban water supply (House 2014). Similarly, there is a wide range of research topics in this set of papers. The most common is an emphasis on sustainability (Chatterley *et al.* 2013, 2014; Kaminsky & Javernick-Will 2014; Welle *et al.* 2015). Interestingly, and as opposed to past trends in the broader sanitation literature (Rosenqvist *et al.* 2016), these papers use sustainability to refer to the sustained use of WASH services with reference to social systems rather than targeting environmental sustainability. Other work studies methods of project delivery such as private participation in WASH infrastructure construction (Santosh Kumar Delhi *et al.* 2012; House 2014) or drivers of conflict regarding these projects (Boudet *et al.* 2011). Several of the papers used QCA in combination with either qualitative or statistical methods for mixed-methods analysis (House 2014; Welle *et al.* 2015). This included one of the few identified publications coming from practice rather than academic researchers (Welle *et al.* 2015). As will be described in more detail below, many of these authors identified an intermediate number of cases and a need for nuanced yet rigorous quantification of complexity as rationales for choosing the QCA method.

QCA VARIANTS

There are three variants of QCA analysis. These variants describe the way that the causal conditions and outcome of interest are defined. In any of these three variants, and if the analytic decisions discussed here are fully documented, QCA enables a fully replicable analysis with transparent measures of reliability. The simplest is *crisp set QCA* (*csQCA*), in which each causal condition is measured as being either fully present or fully absent. For example, Kaminsky & Javernick-Will (2014) use csQCA to describe household toilets that were or were not functional on the day of a research visit. Similarly, Chatterley *et al.* (2013) use csQCA to analyze schools with and without well-maintained toilets. In a slightly more complex variant, *multiple value QCA* (*mvQCA*) permits the inclusion of non-dichotomous measurements which can take on multiple values. mvQCA is best suited for studies in which the variables can be summarized into a small number of discrete options (Gross & Garvin 2011). For example, in a study attempting to understand the effect of water supply operators, we might wish to consider three variants: (1) community-managed supply, (2) private operator, and (3) public operator. An mvQCA study can also include variables that are dichotomized.

Finally, in the most conceptually complex variant, *fuzzy set QCA* (*fsQCA*) allows for each variable to be assigned a value between zero and one corresponding to its degree of membership in a set. In an fsQCA study, a score of 1 represents full membership in a set and a score of 0 represents full non-membership, with 0.5 as the point of maximum ambiguity of set membership. Values between 0 and 1 represent varying degrees of membership and non-membership. These scales are not linear, and the way in which the set is defined will affect the values. fsQCA is useful in cases where restricting all conditions to dichotomous values would result in a loss of detail from data as

it can represent small, but meaningful, differences between cases. For example, in Chatterley *et al.* (2014) the set of schools with well-managed sanitation is defined as schools with toilets that are both functional and clean (Table 2 provides details of Chatterley's definitions). In this paper, schools with excellent performance on both of these metrics are *fully in* the set of schools with well-managed sanitation; schools with moderate performance on these metrics are *partially in* the set of schools with well-managed sanitation; schools with poor performance on these metrics are *fully out* of the set of schools with well-managed sanitation.

THE QCA PROCESS

Defining outcomes and conditions

For all three variants of QCA, the first step is to define the *outcome(s) of interest* to the research question. The outcome of interest is whatever the study intends to measure as an outcome of the intervention; examples in the WASH sector could include open defecation-free status in a community, functionality of a handpump, or household practice of a hygiene behavior. This step informs the process of case selection, as it is necessary to purposefully identify a set of cases that demonstrate a range of the outcomes for the analysis. For example, in Table 1 the hypothetical outcome of interest is Sustained Water Services, meaning cases with and without sustained water service would be needed for the dataset.

Table 1 | Hypothetical example of cases and variables

Outcome of interest: Sustained Water Service			
Observed cases	**Community participation *Condition A***	**Municipal utility *Condition B***	**Set notation[a] representing this combination**
Cases 1–4	Yes	No	A* ~ B
Cases 5–8	No	Yes	~A*B
No observed cases	No	No	~A* ~ B
No observed cases	Yes	Yes	A*B

[a]In set notation, the symbol * signifies and, while ~ signifies not.
Note: hypothetical data.

Table 2 | Excerpted fsQCA coding scheme from Chatterley *et al.* (2014)

Well-managed sanitation services
Minimum of the following two measures: *Reliably functional toilets*[a]:
1: students have reliable access to functional services; repairs timely addressed
0.67: all toilets usually function, but repair needs are not always timely addressed
0.33: some toilets are frequently unusable; repairs are not timely addressed
0: students do not have reliable access; repairs are rarely addressed
and *Reliably clean toilets*[b]:
1: all toilets are almost always clean and quickly cleaned when dirty
0.67: usually more or less clean, with some instances where they remain dirty
0.33: frequently unclean and are usually considered unclean by students
0: rarely clean and students label them as dirty

[a]'Functional' = waste is easily flushed, the building structure, doors and locks function providing privacy, water is available, and soap is available in or near the toilet. 'Repairs timely addressed' = minor critical repairs (needed for use), such as a door lock or clogged toilet, are repaired within 24 hours, major critical repairs, such as a broken pan or door, are repaired within 1 week, minor non-critical repairs, such as a broken tap, are repaired within 1 week, and major non-critical repairs, such as a broken water pump, are repaired within 1 month.
[b]'Clean' = no visible feces on the floor/walls/seat, no flies, and no foul smell.

The next step is to identify *conditions* that are expected to influence the outcome(s) of interest. Conditions are the variables that distinguish one case from another. For example, in Table 1 we provide a hypothetical example where the causal conditions are Community Participation and a Municipal Utility. The selection of conditions for any QCA study is iterative. Conditions are logically constructed and should generally be grounded in theory. However, one of QCA's strengths is the ability to build theory from the analysis. Thus, some of the conditions may be selected for inductive reasons, meaning additional conditions may emerge during the data collection. Indeed, it is likely that a large number of conditions will be identified (Amenta & Poulsen 1994). However, each new condition adds complexity to the logic space (the space defined by all of the possible value-combinations of the conditions (Ragin 2014)), so it is practically important to limit the number of conditions.

A large number of conditions will likely result in a unique explanation for each case, making it difficult to interpret and generalize the results. As such, there are various documented techniques to reduce the causal conditions in a QCA analysis (Ragin 2008; Rihoux & Ragin 2008). For example, it is possible that some of the identified causal conditions will not vary significantly between the cases selected because of the research context. As in statistical analysis, the non-varying conditions cannot be included in the analysis. Such variables are called *domain conditions*. While these domain conditions cannot be analyzed, they may have an important influence on the outcome through their presence and interactions with other causal conditions. It is important to clearly describe any domain conditions, as these limit the generalizability of the results. It may also be possible to combine initially hypothesized conditions if interrelationships between the conditions can be identified. For example, discriminant analysis can be used to identify strong bivariate relationships, or composite conditions can be created through techniques such as factor analysis (Jordan *et al.* 2012).

Case selection

For QCA analysis, cases are selected to exhibit the greatest possible variety of configurations (a *configuration* is defined by each case's set of condition and outcome values). Although many criticize the conscious selection of cases as an improper manipulation of the dataset, this practice is appropriate for QCA because the method's logic is not probabilistic: that is, it does not matter if only a few cases exhibit certain conditions (Berg-Schlosser *et al.* 2009). Rather, the selection of cases exhibiting maximum variation in condition and outcome values will result in the richest possible explanations of relationships among the variables (Gross 2010). The number of cases included in the analysis should be driven by the size of the logic space (the number of all possible configurations) and the feasibility of data collection. A key strength of QCA analysis is that it allows researchers to handle an intermediate number of cases, too many for qualitative analysis and too few for statistical analysis. As we will discuss below, the use of mathematics to search for patterns in case data reduces the data processing demands that logistically limit the scale of traditional qualitative analyses. Similarly, while there is no upper bound to the number of cases QCA can consider, the method does not require minimum sample sizes. This combination means QCA fills the methodological need for a way to rigorously handle an intermediate numbers of cases.

Unobserved configurations are theoretical combinations of conditions that are not found in any of the cases analyzed, and will appear in any QCA study. For example, in the hypothetical dataset represented in Table 1, we do observe cases with the outcome of Sustained Water Service when we observe Community Participation without a Municipal Utility (Cases 1–4), and vice versa (Cases 5–8). However, while they are logically possible we do not observe any cases with neither Community Participation nor a Municipal Utility, or any cases with both Community Participation and a Municipal Utility. These are called unobserved configurations, regardless of the likelihood of their existence. During QCA analysis, unobserved configurations may be handled in three standard ways; the researcher must determine which of these is most appropriate to the research question and data. The different assumptions in each of these three methods should be expected to result in different answers, which are called the complex, intermediate, and parsimonious solutions. These are discussed in more detail in the Pathway analysis section below.

Data collection

Data must be collected for each condition in each case. Generally, data will be collected on more conditions than are actually used in the analysis, given the iterative nature of the QCA process. Both qualitative and quantitative data can be used in the analysis, but the researcher must have sufficient knowledge of each case to make determinations about the variable calibration described below.

Variable calibration

Once the cases, outcomes, and preliminary conditions have been determined, raw data for each case (qualitative and/or quantitative) must be collected and calibrated according to set definitions relevant to the research questions, underpinning theory, and the dataset itself. These set definitions

rigorously describe each case in terms of the causal conditions that are hypothesized to be relevant to the outcome of interest. For example, and as described in more detail below, set definitions might include what kind of management scheme is used for water system management, the volume of water used by each household, or how wealthy communities are.

The method of calibration depends on the variant of QCA being undertaken. For a csQCA study, all conditions must be calibrated as either a zero or one. Qualitative data are calibrated by defining the features of what is within and outside of the set. For example, Kaminsky & Javernick-Will (2014) coded toilets as either socially sustainable (1, defined as owner maintenance post-construction and unbroken slab, pit rings, and superstructure on the day of the visit) or unsustainable (0, if either of the two criteria were not met) on the day of a research visit. Quantitative data are dichotomized through the determination of a numeric cutoff point. For example, water samples might be coded as having either positive or negative fecal coliform test results, based on international standards for water quality. In contrast, for an mvQCA study a small number of discrete options are defined for each condition. For example, community water supply could be coded as community managed (0), having a private operator (1), or having a public operator (2).

Calibration for an fsQCA study can be more complex, as each variable is represented on a continuous scale between zero and one. The first method for calibration, *direct calibration*, can be used for quantitative data. However, quantitative data cannot simply be normalized to values between 0 and 1, as the calibrated values represent the degree of membership in a set and must be based on the set definitions. To perform direct calibration, the research must specify three breakpoint values: full membership, full non-membership and the crossover point (equal to a 0.5 score). These points should be anchored in external criteria and theoretical and case knowledge. For example, in examining country level data for GNP to assess membership in the set of poor countries, the variation between countries that are clearly outside of the set of poor countries is unimportant to the analysis, and the anchor points must be set accordingly. For example, both Sweden and Norway are non-poor, and any variation between the two is unimportant to the set classification of poor countries. As such, the scores for both these nations would be set to 0 (indicating non-membership in the set of poor countries).

The second method for calibrating fsQCA data is *indirect calibration*. This can be done for either quantitative or qualitative data by creating groupings of cases. To calibrate qualitative data the researcher develops a list of operationalized measures for each of the conditions and outcomes. Then, qualitative anchors are defined for full membership and non-membership in the set and the case data are evaluated based on these operationalized definitions. As for direct calibration, these points should be anchored in external criteria and theoretical and case knowledge. For example, Table 2 shows an example of indirect calibration from the literature (Chatterley *et al.* 2014), where schools that virtually always have clean and functional toilets are defined as being fully in the set of schools with well-managed sanitation.

For any of these methods of calibration, a clear calibration protocol and inter-calibrator reliability checks are needed to support the validity of the findings. Through this process, it is likely that the calibration methods will be iteratively improved to ensure that real differences between cases are captured accurately.

Constructing and analyzing the truth table

The calibrated data are used to populate a truth table that represents the calibrated conditions and outcomes. The *truth table* (Table 3) consists of columns for each condition and outcome, with rows representing each case. Once the truth table is generated, the researcher may find *contradictory configurations*, or cases with identical conditions and differing outcomes. These can be resolved by considering the conditions included to see whether (for example) there is a missing condition that explains the difference between the two cases. Alternatively, calibration cutoffs may be re-examined to determine if an important difference between the two cases was obscured in the initial calibration. Researchers intending to use QCA should note that the creation of a contradiction-free truth table is extremely time consuming and requires deep case knowledge.

Table 3 | Truth table example

	Condition A	Condition B	Condition C	Condition D	Outcome
Case 1	1	1	1	0.67	1
Case 2	0.33	0.67	0	1	0
Case 3	0	1	0.33	0	0
…	…	…	…	…	…
Case N	0	0.33	0.67	1	1

Note: hypothetical data.

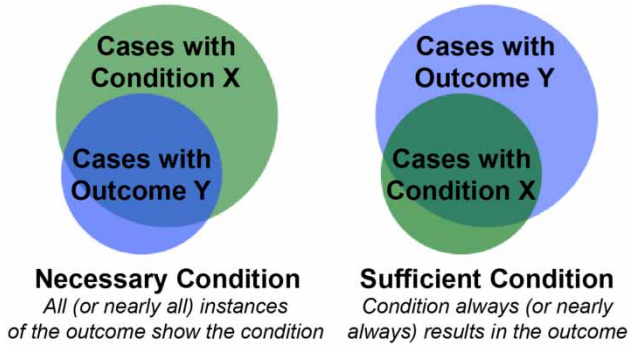

Figure 1 | Necessity and sufficiency.

The first step in QCA analysis determines which individual conditions are necessary or sufficient to achieve the outcome. The second step determines which combinations of conditions combine to lead to the achievement (or non-achievement) of an outcome. These two steps are discussed in more detail below. Although these analyses can be performed by hand, software can facilitate analysis. One option is the open source fs/QCA software developed by Charles Ragin, which can be used for csQCA or fsQCA. Other options are Tosmana, a software package designed for mvQCA and csQCA studies, or various and constantly evolving options for STATA and R (Jordan *et al.* 2012; Schneider & Wageman 2012; Ragin *et al.* 2013)

Necessity and sufficiency of individual conditions

In QCA, *sufficiency* is a measure of the degree to which an individual causal condition is a subset of the outcome (see Figure 1). If a specific condition always (or nearly always) results in a positive outcome, that condition would be deemed sufficient. For example, when we described the intermediate solution for the hypothetical example given in Table 1, the presence of Community Participation was enough to achieve Sustained Water Services, regardless of the presence or absence of a Municipal Utility. Another way to say this is to say that Community Participation is sufficient to achieve Sustained Water Services. In contrast, *necessity* measures the degree to which the outcome is a subset of individual causal conditions, meaning that, if all (or nearly all) cases where the outcome is present have a particular condition present, we would consider that condition necessary. For example, Bogler & Meierhofer (2015) find that both trouble-free production and high demand are necessary for sustainable colloidal silver filter businesses in Nepal.

Mathematically, necessity and sufficiency may be represented by Equations (1) and (2) respectively, where X_i and Y_i represent single conditions. Typically, researchers require a necessity score of at least 0.9 to call a condition necessary for the outcome of interest, and a sufficiency score of at least 0.8 to call a condition sufficient for the

Table 4 | Symmetric and non-symmetric relationships

	Outcome of interest: Unsustained Water Service	Outcome of interest: Sustained Water Service
Symmetric relationship (chi-squared $p < 0.000$)		
Causal condition: Community Participation absent	10	0
Causal condition: Community Participation present	0	10
Non-symmetric relationship (chi-squared $p = 0.12$)		
Causal condition: Community Participation absent	10	40
Causal condition: Community Participation present	0	10

Note: hypothetical data.

outcome of interest.

$$\text{Necessity} = \frac{\sum(\min(X_i,\ Y_i))}{\sum X_i} \quad (1)$$

$$\text{Sufficiency} = \frac{\sum(\min(X_i,\ Y_i))}{\sum Y_i} \quad (2)$$

The analysis of combinations of conditions, a key strength of QCA, is discussed below.

Pathway analysis of combinations of conditions

The truth table is analyzed using Boolean algebra (for csQCA and mvQCA studies) and fuzzy logic (for fsQCA studies). For more details on the mathematics behind these analyses see Ragin (2008, 2014). Regardless of which mathematical approach is used, the analysis of the truth table results in the discovery of combinations of conditions (often called *pathways*) that lead to a particular outcome of interest, with quantitative scores that describe how well each of these pathways describes the dataset. For example, in Table 1 there were multiple cases that showed the conditions of Community Participation, no Municipal Utility, and the outcome of Sustained Water Services; these cases share a pathway.

Two metrics are employed to assess QCA pathway outputs: *consistency* and coverage. Consistency is a measure of the degree to which cases sharing the same combination of conditions have the same outcome. In other words, consistency is a measure of the extent to which the observed cases align with each other. High consistency means a given pathway almost always leads to a certain outcome, while low consistency means a given pathway only sometimes leads to the outcome of interest. Necessity and consistency use the same equation (Equation (1)) but consistency describes a particular combination of conditions rather than considering individual conditions. As such, to measure consistency using Equation (1), X_i represents the membership in a configuration, and Y_i represents the membership in the outcome condition. A consistency score of 1 would indicate perfect consistency, where all cases with a given set of causal conditions have membership in the outcome set to a greater degree than membership in the configuration; however, consistency scores above 0.8 are generally considered acceptable. According to Ragin, 'Consistency, like significance, signals whether an empirical connection merits the close attention of the investigator. If a hypothesized subset relation is not consistent, then the researcher's theory or conjecture is not supported.'

In contrast, *coverage* is a measure of how much a particular pathway accounts for the instance of the outcome, giving a measure of the importance of that pathway. Sufficiency and coverage use the same equation (Equation (2)) but coverage describes a particular combination of conditions rather than considering individual conditions. As such, to measure coverage using Equation (2), X_i represents the membership in a configuration, and Y_i represents the membership in the outcome condition. High coverage scores indicate that a given pathway represents many of the represented cases. However, this does not mean that pathways with low coverage are unimportant, as QCA is not probabilistic. Despite this, knowing which pathways to a given result are seen more frequently can help guide practitioners to interventions that may be more likely to apply to many cases.

Complex, parsimonious, and intermediate solutions

The truth table pathways analysis results in three different solutions: complex, parsimonious, and intermediate (as described below). These different solutions are based on different assumptions made about the unobserved configurations discussed above in the section entitled 'Case selection'. QCA uses *counterfactual analysis* to transparently compare the impacts of assumptions regarding unobserved configurations, and to obtain more parsimonious solutions based on these unobserved configurations. However, it is up to the researcher to use her theoretical and substantive knowledge to decide to what degree these unobserved configurations should be included in the analysis. For example, in the hypothetical example given in Table 1 we might not expect to see cases of Sustained Water Services with the absence of both Community Participation and a Municipal Utility, but we probably would suspect that there are unobserved cases with the presence of both of these conditions. To validate this intuition, published literature from academic journals or practice can be used. Alternatively, more research would be needed to seek out additional cases and better populate the logic space.

As such, there are three possible solutions from each QCA run, depending on how unobserved configurations are used during simplification (ranging from none to all those logically possible). Firstly, the *complex solution* does not incorporate any counterfactuals and is based entirely on the observed cases. This will often be a highly complicated solution, sometimes with a unique pathway for each observed case. It is not typically used by researchers as it does not take into account any theoretical knowledge about the link between conditions and outcomes.

In contrast, the *intermediate solution* uses 'easy' counterfactuals, which are researcher-specified theoretical assumptions. For example, a researcher may believe that three conditions (A, B and C) are likely related to the positive instance of an outcome, but only observes cases where A and B are present, but C is absent (i.e. A*B*~C). If the researcher has strong knowledge that the presence of C should contribute to the outcome under the scenario, then the assumption that A*B*C would lead to the outcome would be an 'easy' counterfactual. It should be noted that these types of assumptions are common, but usually implicit, in traditional comparative case study analysis. A strength of QCA is that these assumptions are clearly documented throughout the analysis process.

Thirdly and finally, the *parsimonious solution* is obtained by using all of the unobserved configurations as potential simplifying assumptions in the truth table analysis. In this solution, the researcher does not specify which assumptions are reasonable, but rather allows the software to find the mathematically simplest solution. Clearly, the researcher must evaluate each of the assumptions that result from the parsimonious solution algorithm to ensure that they are theoretically plausible. It is quite likely that (as in the example just given) at least some of the assumptions leading to the parsimonious solution would be difficult to justify. As such, the intermediate solution should be reported unless there are strong theoretical reasons to accept the parsimonious solution.

Note that it is also recommended to perform an analysis of which conditions lead to the lack of attainment of an outcome. Because QCA accounts for configurational complexity and asymmetrical relationships (which are discussed in more detail below), this will not necessarily be the negation of the conditions that led to the outcome. This asymmetry exists because the question of what conditions lead to a positive outcome is not necessarily the same question as which conditions lead to the negation of that outcome. Therefore, this should be treated as a separate truth table analysis, and interpreted using the same procedures as the analysis for the positive outcome. For example, Chatterley *et al.* (2013) report pathways to both well-maintained school toilets (the positive outcome) and pathways to poorly maintained school toilets (the negative outcome). In that study, poor construction is a causal condition that appears in all pathways to poorly maintained school toilets, while its inverse (quality construction) is a condition in only some of the pathways to well-maintained school toilets. In other words, poor quality construction is present in all cases with the poorly maintained school toilet outcome. However, good quality construction is not present in all cases with well-maintained school toilets. This suggests poor construction can be overcome, given the presence of a number of other conditions such as (for example) a local sanitation champion.

CRITIQUES OF QCA

As for any method, there have been important critiques made of QCA. Most recently, there has been a flurry of research attention that uses simulations to examine the robustness of QCA findings. For example, Hug (2013) claims that QCA does not allow researchers to directly account for measurement error and uses a quasi-Monte Carlo analysis to demonstrate the implications of this for research conclusions. In another example, Braumoeller (2015) makes the strong claim that QCA does not permit researchers to discover whether or not their findings are the result of chance, noting that QCA does not generate statistical significance tests. Similarly, Krogslund *et al.* (2015) note that the results of QCA are sensitive to researcher decisions such as cutoffs for the minimum frequency of cases that are included in analysis and the minimum and maximum sufficiency scores required for the analysis. More troublingly, and related to Braumoeller's critique, they also claim that QCA suffers from confirmation bias. As might be expected, these various critiques have been answered by other QCA simulation analyses that claim

(for example) that many of the issues observed in the critical simulations stem from a fundamental misunderstanding of the QCA method and analysis procedures (for one example, see Rohlfing 2015).

It is not our purpose here to undertake quantitative modeling to contribute to this scholarly conversation. Instead, we would note that similar critiques could be leveled at most research methods. Any analysis can be undermined by undetected measurement errors; QCA's dependence on deep, qualitative case knowledge is an important answer to this problem and is one of the key strengths of the method. It is also true that QCA does not generate measures of statistical significance. However, there are links between statistical significance and the quantitative consistency values generated by QCA, as discussed by Ragin (2000). The required values for consistency, sufficiency, etc. are indeed researcher selected in QCA, much as the required minimum p value for statistical significance is researcher selected in regression analysis. However, and once again paralleling good research practice in statistics, past research establishes guidance for what acceptable values for these cutoffs are and what the risks of deviating from these standards are.

An equally important critique coming from qualitative research traditions would emphasize that the calibration scales required for QCA analysis are deeply – and potentially problematically – simplified representations of reality. As such, there is a risk that the numeric scores provide a sense of false precision. In this sense, QCA may be seen as overly positivistic and reductive. In response to these important critiques, we acknowledge that the use of any particular research method is extremely unlikely to enable perfect project outcomes (as defined by either the development community or as defined by the end users). However, we do argue that methodological diversity can help us move towards more sustainable WASH infrastructure, by which we mean infrastructure that is used and maintained by communities over time.

WHY AND WHEN TO USE QCA

The preceding sections described how to perform a QCA analysis; the following sections outline why and when QCA is appropriate. Generally, QCA provides a middle ground between traditional case study methods and large n statistical methods, enabling cross-case comparisons to identify patterns across cases while retaining sensitivity to contextual detail. It relies on set-theoretical descriptions of cases, and can be used for exploratory analysis intended to build theory or to test existing theories. Many qualitative researchers already implicitly use set theory to describe findings of case studies by examining cases that share an outcome and determining what causal conditions they share; QCA provides a numeric approach to discovering and documenting these relationships, with transparent documentation of each step in the analysis. For a related discussion of common pitfalls in the use of QCA, we refer the reader to Jordan *et al.* (2012).

For WASH research, an important attribute of QCA is its ability to handle *combinations of qualitative and quantitative data*. For example, in their fsQCA analysis of the challenges facing the production and marketing of colloidal silver water filters in Nepal, Bogler & Meierhofer (2015) were able to consider quantitative data, like population density and percentage of people treating water, alongside qualitative data, such as the reasons customers gave for not buying filters and strategies used for new customers.

A key conceptual difference between QCA and more traditional statistical methods is the different assumption regarding symmetry of relationships. Table 4 uses hypothetical data and a highly simplified example to describe why this assumption is important. The uppermost portion of the table shows an example of a perfectly symmetric relationship. Here, we see that when Community Participation is present, we achieve the outcome of Sustained Water Services. In contrast, when Community Participation is absent, we see abandoned water systems. Given these data, both statistical analysis and QCA find a strong link between Community Participation and Sustained Water Services. However, it would also be possible to find an *asymmetric relationship* between these variables, as shown in the lower portion of Table 4. Here, we see that all cases with Community Participation achieve the Sustained Water Services outcome. However, when Community Participation is absent some cases experienced Sustained Water Services and others showed Unsustained Water Services. Using statistical analysis techniques such as a chi squared test, we would believe that Community Participation is not a statistically significant

factor in Sustained Water Services; the p value resulting from a chi squared test is 0.12, which is above typical cutoffs for statistical significance. In contrast, in set-based analysis like QCA, we would describe these as set relationships and gain the insight that while Community Participation does seem to be sufficient for achieving Sustained Water Services, it is not necessary for achieving it. In other words, Community Participation is a valuable strategy, but not the only one.

The example in Table 4 also shows the importance of considering *configurational complexity* in socially influenced research topics like WASH. While regression methods examine the relative contribution of variables, holding other modeled variables equal, QCA seeks combinations of variables that lead to an outcome and recognizes that there are likely several different combinations of factors that may result in an outcome. This allows us to handle situations where uniformity of causal effects cannot be assumed. For example, in the hypothetical example shown in Table 4 we might add another causal condition such as the presence of a municipal utility (as we did earlier in Table 1). This resolves the excerpted hypothetical data in Table 4 showing that to achieve Sustained Water Services a community requires some combination of Community Participation, a Municipal Utility, or both working together. This hypothetical finding means that Community Participation and the presence of a Municipal Utility are *substitutable conditions*, or conditions that are interchangeable in terms of achieving the outcome of Sustained Water Service.

CONCLUSION

QCA has not been used frequently in studies of water, sanitation and hygiene interventions to date. However, given its ability to account for configurational complexity through the use of Boolean algebra or fuzzy logic, it is a promising method for WASH researchers that strongly complements more commonly used methods. QCA is a useful method to apply when there is a desire to examine the complex interactions between conditions that may influence the success of interventions, and particularly when there are strong reasons for considering both quantitative and qualitative data. In addition, QCA is useful in complex situations where researchers or practitioners are interested in learning if there is more than one way to attain an outcome of interest. Similarly, QCA requires explicit calibration definitions for both conditions and outcomes; this ability to deal with and document nuanced differences is a key strength of the method. In addition, the process of condition calibration is a systematic way for researchers to incorporate unexpected complexities and nuances that emerge during the analysis. As such, QCA is well-suited for examining not only if a particular intervention results in the outcome we expect, but how and why such an intervention does (or does not) work. Results of QCA studies include analysis of what variables are necessary and/or sufficient to achieve an outcome and pathways demonstrating what possible combinations of variables may lead to an outcome. To enable reproducibility and transparency of analysis, documentation of the various analytic decisions detailed in this paper should be reported for every QCA analysis. To date, the majority of QCA studies have been from academics, but QCA has strong potential to be useful to practitioners as well. Often, evaluations of WASH projects rely on qualitative methods; QCA offers a complementary approach to practitioners who wish to rigorously evaluate the success of their projects across a limited number of cases in order to gain an understanding of why interventions may have succeeded or failed in different contexts.

ACKNOWLEDGEMENTS

The authors are grateful to Dr Rachel Peletz and the journal's anonymous reviewers for their comments on drafts of this manuscript.

REFERENCES

Amenta, E. & Poulsen, J. D. 1994 Where to begin a survey of five approaches to selecting independent variables for qualitative comparative analysis. *Sociological Methods & Research* **23** (1), 22–53.

Annamalai, T. R., Narayanan, S., Devkar, G., Kumar, V. S., Devaraj, R. & Mahalingam, A. 2016 *Protocol for a Systematic Portfolio Review: To What Extent Have the Water, Sanitation, and Hygiene (WASH) Sub-Sectors Incorporated the Life-Cycle Approach into Policies, Programmes and*

Projects during Teh MDG Period? New Delhi, India: International Initiative for Impact Evaluation. Retrieved from http://www.3ieimpact.org/en/evidence/systematic-reviews/details/337/.

Berg-Schlosser, D., De Meur, G., Rihoux, B. & Ragin, C. C. 2009 Qualitative comparative analysis (QCA) as an approach. In: *Configurational Comparative Methods*, vol. 51. Sage Publications, Los Angeles, pp. 1–18.

Bogler, A. & Meierhofer, R. 2015 The challenge of producing and marketing colloidal silver water filters in Nepal. *Water* **7** (7), 3599–3612.

Boudet, H., Jayasundera, D. & Davis, J. 2011 Drivers of conflict in developing country infrastructure projects: experience from the water and pipeline sectors. *Journal of Construction Engineering and Management* **137** (7), 498–511. http://doi.org/10.1061/(ASCE)CO.1943-7862.0000333.

Braumoeller, B. F. 2015 Guarding against false positives in qualitative comparative analysis. *Political Analysis* **23** (4), 471–487. http://doi.org/10.1093/pan/mpv017.

Chatterley, C., Linden, K. G. & Javernick-Will, A. 2013 Identifying pathways to continued maintenance of school sanitation in Belize. *Journal of Water Sanitation and Hygiene for Development* **3** (3), 411–422.

Chatterley, C., Javernick-Will, A., Linden, K. G., Alam, K., Bottinelli, L. & Venkatesh, M. 2014 A qualitative comparative analysis of well-managed school sanitation in Bangladesh. *BMC Public Health* **14** (1), 1.

Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., Ensink, J., Freeman, M., Jenkinds, M., Odagiri, M., Ray, S., Sinha, A., Suar, M. & Schmidt, W. 2014 Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health* **2** (11), e645–e653.

Gross, M. 2010 *Aligning Public-Private Partnership Contracts with Public Objectives for Transportation Infrastructure*. Virginia Polytechnic Institute and State University. Retrieved from http://vtechworks.lib.vt.edu/handle/10919/28785.

Gross, M. & Garvin, M. 2011 Structuring PPP toll-road contracts to achieve public pricing objectives. *Engineering Project Organization Journal* **1** (2), 143–156. http://doi.org/10.1080/21573727.2011.572256.

Guiteras, R., Levinsohn, J. & Mobarak, A. M. 2015 Encouraging sanitation investment in the developing world: a cluster-randomized trial. *Science* **348** (6237), 903–906.

House, R. S. 2014 *Public Utilities in the Age of Partnership: Lessons from Private Participation in Urban Water Supply*. National University of Singapore, Singapore. Retrieved from http://scholarbank.nus.edu.sg/bitstream/handle/10635/113292/RSH_WaterPPP_Dissertation_Nov2014.pdf?sequence=1.

Hug, S. 2013 Qualitative comparative analysis: how inductive use and measurement error lead to problematic inference. *Political Analysis* **21** (2), 252–265. http://doi.org/10.1093/pan/mps061.

Jordan, E., Gross, M., Javernick-Will, A. & Garvin, M. 2011 Use and misuse of qualitative comparative analysis. *Construction Management & Economics* **29**, 1159–1173. http://dx.doi.org/10.1080/01446193.2011.640339.

Kaminsky, J. & Javernick-Will, A. 2014 The internal social sustainability of sanitation infrastructure: a set theoretic approach. *Environmental Science & Technology* **48** (17), 10028–10035. http://doi.org/10.1021/es501608p.

Krogslund, C., Choi, D. D. & Poertner, M. 2015 Fuzzy sets on shaky ground: parameter sensitivity and confirmation bias in fsQCA. *Political Analysis* **23** (1), 21–41. http://doi.org/10.1093/pan/mpu016.

Meyer, J. W. & Rowan, B. 1977 Institutionalized organizations: formal structure as myth and ceremony. *American Journal of Sociology* **83** (2), 340–363.

OpenIDEO 2016 *Appraisal Tools for Sanitation Micro Finance*. Retrieved October 3, 2016, from https://challenges.openideo.com/challenge/water-and-sanitation/second-evaluation/appraisal-tools-for-sanitation-micro-finance.

Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceno, B., Ganguly, S., Colford Jr, J. M. & Gertler, P. J. 2014 The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Medicine* **11** (8), e1001709.

Pickering, A. J., Djebbari, H., Lopez, C., Coulibaly, M. & Alzua, M. L. 2015 Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial. *The Lancet Global Health* **3** (11), e701–e711.

Ragin, C. 2000 *Fuzzy-Set Social Science*. University of Chicago Press, Chicago.

Ragin, C. 2008 *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University Of Chicago Press, Chicago.

Ragin, C. 2014 *The Comparative Method: Moving beyond Qualitative and Quantitative Methods*, 2nd edn. University of California Press, Oakland, CA.

Ragin, C., Drass, K. & Davey, S. 2013 *Fuzzy-Set/Qualitative Comparative Analysis (Version 2.0)*. Department of Sociology, University of Arizona, Tuscon, Arizona. Retrieved from www.fsqca.com.

Rihoux, D. B. & Ragin, C. 2008 *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Sage Publications, Los Angeles.

Rohlfing, I. 2015 Mind the gap: A review of simulation designs for Qualitative Comparative Analysis. *Research & Politics* **2** (4), 1–4. http://doi.org/10.1177/2053168015623562.

Rosenqvist, T., Mitchell, C. & Willetts, J. 2016 A short history of how we think and talk about sanitation services and why it matters. *Journal of Water, Sanitation and Hygiene for Development* **6** (2), 298–312. http://doi.org/10.2166/washdev.2016.118.

Santosh Kumar Delhi, V., Mahalingam, A. & Palukuri, S. 2012 Governance issues in BOT based PPP infrastructure projects in India. *Built Environment Project and Asset Management* **2** (2), 234–249.

Schmidt, W.-P. 2014 The elusive effect of water and sanitation on the global burden of disease. *Tropical Medicine & International Health* **19** (5), 522–527. http://doi.org/10.1111/tmi.12286.

Schneider, C. Q. & Wageman, C. 2012 *How To … Use Software for Set-Theoretic Analysis: Online Appendix to 'Set-Theoretic Methods for the Social Sciences'*. Cambridge University Press, Cambridge, MA. Retrieved from www.cambridge.org/US/download_file/134448/.

Spencer, J. H. 2008 Household strategies for securing clean water the demand for piped water in Vietnam's peri-urban settlements. *Journal of Planning Education and Research* **28** (2), 213–224.

Welle, K., Williams, J., Pearce, J. & Befani, B. 2015 *Testing the Waters: A Qualitative Comparative Analysis of the Factors Affecting Success in Rendering Water Services Sustainable Based on ICT Reporting*. WaterAID, Uganda. Retrieved from http://opendocs.ids.ac.uk/opendocs/handle/123456789/7099.

First received 3 October 2016; accepted in revised form 18 January 2017. Available online 8 March 2017