

Review Paper

Improving the use of qualitative comparative analysis for inferring complex causation in development and planning research

Alrik Thiem

ABSTRACT

Empirical research methods provide the necessary means to extract relevant information from data. Qualitative Comparative Analysis (QCA), one such method, is currently making first inroads into the development and planning (D&P) community. On the one hand, QCA is well suited for building empirically founded theories emphasizing causal complexity. On the other hand, however, current use of QCA in D&P research is marked by problematic applications of this method whose results misrepresent the empirical evidence marshaled to support them. Policy recommendations that stand on shaky grounds have been issued in consequence. By reanalyzing a recent empirical study on school sanitation maintenance in Belize, this method workshop article shows how the use of QCA can be improved, which should in turn lead to more solid, evidence-based policy recommendations for development interventions.

Key words | causal complexity, configurational comparative methods, Qualitative Comparative Analysis (QCA), research methods

Alrik Thiem
University of Lucerne,
Lucerne,
Switzerland
E-mail: alrik.thiem@unilu.ch

INTRODUCTION

Empirical research methods provide the means to extract useful information from data. One such method, which is currently making first inroads into the development and planning (D&P) community, is Qualitative Comparative Analysis (QCA) – a so-called *configurational comparative method* of causal inference pioneered by US sociologist Charles Ragin (1987). So far, applications of QCA in D&P research include, for instance, the study of agricultural irrigation systems (Lam & Ostrom 2010), female enrollment in primary education (Holvoet & Inberg 2016), maintenance of sanitation infrastructure (Chatterley *et al.* 2013; Hacker & Kaminsky 2017), and water services management and sustainability (Gasparro & Walters 2017; Person *et al.* 2017). Moreover, methodologically oriented articles have argued for the immense potential of QCA for addressing the complexity inherent in D&P research projects (e.g., Bradshaw *et al.* 1995; Kaminsky & Jordan 2017).

On the one hand, QCA is well suited for building empirically founded theories that emphasize causal complexity. On the other hand, however, current use of this method in D&P research is marked by many problematic applications whose results misrepresent the empirical evidence marshaled to support them. Policy recommendations that stand on shaky grounds have been issued in consequence. By reanalyzing the recent empirical study by Chatterley *et al.* (2013) on school sanitation maintenance in Belize, this method workshop article illustrates how the use of QCA can be improved in inferring complex causation, which in turn should lead to more solid, evidence-based policy recommendations for development interventions.

The article is structured as follows. First, the theory of causation that underlies QCA is introduced. In this connection, it is also explained how QCA operationalizes this theory. Subsequently, the study by Chatterley *et al.* (2013)

will serve as a real-world example whereby the consequences of three methodological problems often affecting current use of QCA in D&P research are illustrated. An annotated replication file for the QCApro extension package for the R environment (Thiem 2018), whereby all steps carried out herein can be examined, repeated, or modified, is available online. The conclusions of this article emphasize that QCA possesses great potential for D&P research, but incorrect use may lead to unsupported policy recommendations and, ultimately, ineffective or even counterproductive development interventions.

WHAT QCA IS AND WHAT IT DOES

QCA is a configurational comparative method of causal data analysis that seeks to identify causal pathways to an outcome of interest (Rihoux & Ragin 2009; Thiem *et al.* 2016). Kaminsky & Jordan (2017: p. 197) rightly note in this connection that QCA's 'causality claim [...] is founded on theory', but the authors do not specify which theory that is, let alone what it consists in. However, such information is crucial for understanding the assumptions, mechanics, and output of QCA. Among the different theories of causation discussed and developed by philosophers of science over the last centuries (Beebe *et al.* 2009), one particular theory provides the epistemological foundations of QCA.

This theory is John L. Mackie's INUS theory (Mackie 1974), a member of the family of regularity theories of causation (Graßhoff & May 2001; Cartwright 2007: pp. 34–35; Baumgartner 2008; Ragin & Strand 2008: pp. 431–432; Psillos 2009; Reiss 2009: p. 23; Schneider & Wagemann 2012: p. 79; Thiem & Baumgartner 2016a: p. 803, 2016b: pp. 347–348). Although its influence has always been localized, the notion of INUS causation has continuously been invoked in numerous sub-disciplines of the natural as well as the social sciences, from economics over medicine and neurology to psychology (e.g., Wulff 1984; Hoover 1990; Meehl 2001; Sawcer *et al.* 2014).

The achievement of Charles Ragin and Kriss Drass, who first pioneered QCA in the mid to late 1980s (Ragin *et al.* 1984; Drass & Spencer 1987; Ragin 1987, 1989), lay in bringing together Mackie's INUS theory with the Quine–McCluskey algorithm (QMC) from electrical engineering

to form the method of QCA. This synthesis triggered nothing less than a methodological revolution (Vaisey 2009). Virtually overnight, and by mere happenstance rather than by intention, it not only solved one of the most vexing problems that had plagued Mackie's theory – the so-called Manchester Hooters Problem – but it also provided a genuine alternative to all other contemporary methods of causal data analysis (e.g., Gerring 2001: p. 209; Reiss 2009: p. 23). Yet, as we will see later, the import of QMC from electrical engineering into social research for purposes of causal inference created new, yet no less serious problems of its own, some of which continue to grip QCA-based empirical work that does not have its finger on the pulse of advanced methodological research on configurational comparative methods.

Because of the fact that the INUS theory remains unfamiliar to most researchers, let us first clarify what its gist is. Suppose, for example, that an economist argues, on the basis of a set of data they had gathered and analyzed, that the severe restriction of credit is a cause of high unemployment (*HUE*) in free market economies (cf. Mackie 1965: p. 252). How is this claim to be interpreted? Undoubtedly, this economist implies neither that severe credit restriction is necessary for *HUE* nor that it is sufficient, let alone both necessary and sufficient. Clearly, other conditions must obtain along with credit restriction for *HUE* to result, including the absence of high government spending (*HGS*), which would offset the unemployment-inducing effect of credit restriction.

Also, this economist is certainly well aware of the fact that unemployment can be brought about by other combinations of conditions in which the restriction of credit plays no role. For example, the introduction of an innovative technology or the relocation of a major company to a country with lower wages may also lead to *HUE* if the laid-off workers do not possess skills that make them immediately employable in sectors other than those replaced by the new technology (*NT*) or lost by that company's relocation, irrespective of whether credit is currently restricted or not. Therefore, the claim this economist is making is merely that the restriction of credit is, in economies of the specified kind, by itself insufficient for *HUE*, yet a non-redundant part of one particular combination of conditions that, although itself unnecessary for *HUE*, is sufficient for it. In short, credit restriction is an INUS condition for *HUE*.

Propositional logic is that branch of the mathematical system of Boolean algebra which operationalizes the concepts of ‘sufficiency’ and ‘necessity’ functionally. Under this system, the economist’s argument about the causes of *HUE* laid out above is syntactically codified as shown in Expression (1):

$$\overbrace{SRC \wedge \neg HGS \wedge X_1}^1 \vee \overbrace{NT \wedge X_2}^2 \vee \overbrace{MCR \wedge X_3}^3 \vee Y \leftrightarrow HUE, \quad (1)$$

where *SRC* denotes ‘severe restriction of credit’, *HGS* ‘high government spending’, *NT* ‘new technology’, *MCR* ‘major company relocation’ and *HUE* ‘high unemployment’, while ‘ \wedge ’ stands for the logical concept ‘and’, also called conjunction, ‘ \neg ’ for the logical concept ‘not’, also called negation, ‘ \vee ’ for the logical concept ‘or’, also called disjunction, and ‘ \leftrightarrow ’ for the logical concept ‘is necessary and sufficient for’/‘if, and only if, then’, also called equivalence (if it was the case that not all instances of *HUE* in a given set of data could be explained, then the equivalence operator would be replaced with the less strong implication operator, ‘ \rightarrow ’, indicating that the disjunction of conjunctions was not necessary for *HUE*). The generic variables X_1 , X_2 , and X_3 stand for further INUS conditions that are causally relevant to *HUE*, but that could not be identified on the basis of the economist’s data, while Y symbolizes the disjunction of all other combinations of INUS conditions that are causally relevant to *HUE* but that could not be identified on the basis of the economist’s data, either. (In this connection, it is important to point out that just because some factor does not appear in a final model, this does not mean that this factor is generally causally irrelevant to the analyzed outcome. It only means that, conditional on the available data, there exists no evidence for the causal relevance of that factor. In the extreme, only very few parts of a causal structure will be discoverable. The quality and quantity of available data will have a major impact on the extent to which an INUS structure is identifiable by means of QCA.)

The analytical challenge QCA faces is to discover INUS conditions and their exact causal interplay, represented by a functional statement such as Expression (1), from a set of

data that has been collected. This discovery process is carried out by optimization algorithms such as QMC (McCluskey 1956). This article is not the place to go into the details of this algorithm, which have been laid out numerous times in the electrical engineering and logic design literature (e.g., McCluskey 1965: pp. 140–157; Edwards 1973: pp. 98–108; Lewin & Protheroe 1992: pp. 76–86; Roth & Kinney 2014: pp. 167–192), but its central principle shall be introduced nonetheless because it is as simple as it is elegant in helping researchers discover INUS conditions and packages of INUS conditions that form alternative causal routes to the outcome. The following passage focuses on path 1 in Expression (1).

Imagine the economist’s data contained a case c_1 , which has experienced severe restriction of credit, the absence of *HGS* as well as *HUE*. Concretely, c_1 is characterized by $SRC \wedge \neg HGS \wedge \neg NT \wedge \neg MCR \wedge HUE$. In the economist’s data are also three further cases of the following kind: c_2 : $SRC \wedge \neg HGS \wedge NT \wedge \neg MCR \wedge HUE$; c_3 : $SRC \wedge \neg HGS \wedge NT \wedge MCR \wedge HUE$; and c_4 : $SRC \wedge \neg HGS \wedge \neg NT \wedge MCR \wedge HUE$. Based on the Boolean-algebraic laws of distribution, complementarity, and identity, QMC iteratively eliminates factors to which causal relevance cannot be attributed in the context of the remaining factors. First, it combines c_1 and c_2 . Because the only exogenous factor that differs across these two cases is *NT*, while the outcome is the same, *NT* is redundant and can be eliminated to yield $SRC \wedge \neg HGS \wedge \neg MCR \wedge HUE$. Second, it would combine c_3 and c_4 . Because the only exogenous factor that differs across these two cases is again *NT*, while the outcome is the same, *NT* is redundant and can be eliminated to yield $SRC \wedge \neg HGS \wedge MCR \wedge HUE$. Such shortened expressions from which at least one factor has been eliminated are called implicants. The two implicants that have resulted from the previous two eliminations can themselves be combined to yield $SRC \wedge \neg HGS \wedge HUE$. If now all remaining cases of *SRC* or $\neg HGS$ in the data are associated with $\neg HUE$, no further eliminations are possible, and QMC concludes that $SRC \wedge \neg HGS$ is a prime implicant, which is the technical term for a potentially causal path in QCA.

It is relatively easy to show that the vast majority of problems in methodological and applied QCA work have their roots in misinterpretations or misapplications of one or more of the elements of the INUS theory or of QMC. The

next section will address the three most serious issues: first, the testing of isolated necessary conditions and their subsequent insertion into QCA solutions; second, the use of solution types that are unsuitable for causal inference under the INUS theory; and third, the elimination of models from consideration that would have provided alternative explanations of the data due to a misapplication of QMC. The recent study by Chatterley *et al.* (2013) will be thoroughly re-analyzed for this purpose.

REVISITING THE CAUSES OF WASH INTERVENTION MAINTENANCE

Water, Sanitation and Hygiene (WaSH) interventions in schools located in low-income countries have received much attention from D&P researchers and practitioners, not least because the success of such interventions is imperative from both an economic as well as a humanitarian perspective. Students with access to functioning and clean facilities benefit not only in terms of their health but also with regard to learning outcomes and educational opportunities (e.g., Chatterley *et al.* 2013; Chatterley *et al.* 2014; Kaminsky & Javernick-Will 2014).

In a recent QCA study and contribution to this body of research, Chatterley *et al.* (2013) sought to identify the combinations of conditions that are conducive to post-implementation maintenance of school-based WaSH interventions. To this end, the authors chose the state of school toilets at 15 case schools in Belize as an indicator of maintenance. A school's sanitation infrastructure was either well-maintained (*WM*) or poorly maintained (*PM*). Six exogenous factors were included to explain the state of maintenance, four of them addressing social aspects of the intervention, and two relating to technical aspects. Factors in the first group were upfront local involvement in program implementation (*LI*), financial or in-kind community support for operation and maintenance of WaSH services (*CS*), the presence of a local champion (*LC*), and the absence of vandalism (*NV*). Factors in the second group were quality construction (*QC*) and familiarity with the implemented sanitation technology (*FT*). The complete data set is provided in Table 1.

Table 1 | Data on state of maintenance of school toilets, adopted from Chatterley *et al.* (2013: p. 416)

School	LI	CS	LC	NV	QC	FT	WM	PM
1	1	0	1	1	1	0	1	0
2	1	1	0	1	1	0	1	0
3	1	1	1	1	0	1	1	0
4	1	1	1	1	1	0	1	0
5	1	1	1	0	1	0	1	0
6	1	1	1	0	1	1	1	0
7	1	0	1	0	1	1	1	0
8	1	0	1	1	1	1	1	0
9	0	0	0	0	0	1	0	1
10	0	0	1	0	0	0	0	1
11	0	0	0	1	0	0	0	1
12	1	0	0	0	0	0	0	1
13	0	1	0	0	0	1	0	1
14	0	0	1	0	0	1	0	1
15	0	1	0	1	0	1	0	1

LI, local involvement; *CS*, community support; *LC*, local champion; *NV*, no vandalism; *QC*, quality construction; *FT*, familiar technology; *WM*, well-maintained school sanitation; *PM*, poorly maintained school sanitation.

Testing for and inserting necessary conditions into QCA models

In line with methodological guidelines formulated in Ragin (2000: p. 106), the first task performed by Chatterley and her co-authors was the search for simple necessary conditions for well-maintained school sanitation and poorly maintained school sanitation, respectively. The authors identified upfront local involvement in program implementation as a perfectly necessary condition for the former and the absence of quality construction as a perfectly necessary condition for the latter. There was not a single case in the data where the outcome was present but the respective condition was not.

Subsequently, the solutions for each outcome were generated using the fs/QCA 2.5 software developed by Ragin & Davey (2014), and the aforementioned necessary conditions were inserted into the models if they were not present. One model consisting of five alternative pathways accounted for well-maintained school sanitation and a model consisting of three alternative pathways for poorly maintained school sanitation. Both models are restated in m_{wm} in Expression

(2) and m_{pm} in Expression (3), respectively, with the leading top-level condition being the inserted, perfectly necessary condition, and lower-case letters indicating the negation of a condition:

$$m_{wm}: LI \wedge \overbrace{(LC \wedge NV \wedge FT)}^1 \vee \overbrace{QC \wedge NV \wedge CS}^2 \vee \overbrace{QC \wedge LC \wedge CS}^3 \\ \vee \overbrace{QC \wedge LC \wedge NV}^4 \vee \overbrace{QC \wedge LC \wedge FT}^5 \leftrightarrow WM \quad (2)$$

$$m_{pm}: qc \wedge \overbrace{(li \wedge lc)}^1 \vee \overbrace{cs \wedge nv \wedge li}^2 \vee \overbrace{cs \wedge nv \wedge ft \wedge lc}^3 \leftrightarrow PM \quad (3)$$

Recall that QCA's search target is a set of causally interpretable models, that is, models that only contain INUS conditions and must, therefore, be free of redundancies. However, neither LI nor qc survived the process of redundancy elimination via QMC in fs/QCA 2.5, which means that neither could be demonstrated to be an INUS condition on the basis of the available data. Why did the authors include these two conditions nonetheless? After all, the inclusion of conditions for which no evidence of causal relevance can be found in the data into otherwise causally interpretable models misrepresents the empirical evidence.

The explanation is that the concept of 'necessity' has been overinflated in much of the QCA literature; that it has been infused with a notion of causality it simply does not have, either in the theory underlying QCA, or in the algorithms or their output that operationalize this theory. Examples from the literature that corroborate this hypothesis are not difficult to find (Thiem & Baumgartner 2016a).

For instance, Mahoney *et al.* (2009: p. 118) argued that in QCA and other related methods 'a necessary cause can be defined as: X_1 is a necessary cause of Y_1 if Y_1 is a subset of X_1 ', that is, if it is true that X_1 is a necessary condition for Y_1 , or for short, $X_1 \leftarrow Y_1$. Similarly, Ragin (2008: p. 53) claimed that '[a]n argument of causal necessity is supported when it can be demonstrated that instances of an outcome constitute a subset of instances of a causal condition'.

Yet, the formal (Boolean) definition of the logical concept of 'necessity' does not entail that X_1 's status as a

superset of Y_1 renders the former a *cause* of the latter (Thiem *et al.* 2016; Thiem 2016, 2017; Thiem & Baumgartner 2016a), and no modern regularity theory of causality argues otherwise (see Beebe *et al.* 2009). This does not mean that necessary conditions cannot be INUS conditions. Indeed, Mackie (1965: p. 246) had already clarified that 'it is not part of the definition of an INUS condition that it should *not* be necessary'. Sometimes, INUS conditions are also necessary conditions. However, the testing for simple necessary conditions in isolation and their insertion into final models in the absence of tests for their causal relevance as INUS conditions eliminates all the progress regularity theories of causation have made over two centuries.

What does this mean for applied research that issues policy recommendations? If researchers emphasize in reports to project sponsors or donors that a particular condition has been found to be necessary for ensuring the success of an intervention, even though it could not be shown to even potentially have an effect on the outcome, scarce resources are likely to be wasted. These resources are not available anymore and thus cannot be spent on creating conditions for which there exists evidence that they help ensure the success of an intervention. The opportunity costs following from this ill-application of QCA may even outweigh the actual costs of the inefficacious intervention itself.

Using solution types that are unsuitable for causal inference

Currently, QCA offers three solution types: conservative solutions (Ragin 1987: pp. 104–113), intermediate solutions (Ragin & Sonnett 2005), and parsimonious solutions (Ragin 1987: pp. 104–113). That these three solution types exist at all is due to a particular feature of QMC, namely, its peculiar approach to identifying prime implicants not via situations of difference on the outcome, but ones of agreement, as was shown above. Had Ragin imported not QMC, but another minimization algorithm, the QCA literature that has evolved over the past two decades might have looked very different. However, Ragin's decision is not surprising, given that QMC is *the* standard algorithm described in every introductory textbook on electrical engineering or logic design of the last five decades (e.g., McCluskey 1965: pp. 140–157;

Edwards 1973: pp. 98–108; Lewin & Protheroe 1992: pp. 76–86; Roth & Kinney 2014: pp. 167–192).

The feature of QMC that has worried social scientists ever since the publication of Ragin (1987) is its use of so-called *don't cares*, that is, conjunctions of conditions that do not exist in the data for whatever reason, either because the analyst ran out of resources, because nature has not (yet) furnished researchers with the possibility to observe such conjunctions, or because a common-cause structure lies behind the data. In QCA, these conjunctions are called *logical remainders*. If QMC finds one of these remainders to differ from an empirically instantiated conjunction in one single factor only, the algorithm declares the former to be sufficient for the outcome in order to be able to draw on the Boolean-algebraic laws of distribution and complementarity for eliminating that factor.

However, declaring some remainder *A* to be sufficient for some outcome *B* has been interpreted by many QCA researchers to mean that the existence of *A* would be (counterfactually) assumed, and assumed so in combination with the existence of *B*. Ragin thus proposed conservative and intermediate solutions to block QMC's access to remainders by declaring some or all of them not to be sufficient for the outcome. Following recommendations issued by Ragin (2008: p. 171), Kaminsky & Jordan (2017: p. 204) thus tell WaSH researchers that intermediate solutions are optimal and should be reported because they block QMC from using so-called difficult counterfactuals.

These recommendations are based on erroneous assumptions because the definition of an implication, which underlies every statement of sufficiency, also specifies that the falsity or non-existence of *A* makes the implication true (the proposition $A \rightarrow B$ is true for $\{A = 0, B = 1\}$, $\{A = 0, B = 0\}$, and $\{A = 1, B = 1\}$). Hence, the use of logical remainders by QMC is unproblematic from the perspective of causal inference with empirical data, not least because no QCA solution will ever feature a causal path that has been derived on the basis of logical remainders alone. The use of logical remainders is a mere algebraic shortcut on QMC's way to output minimally necessary disjunctions of minimally sufficient conjunctions, which are functional statements that are causally interpretable in accordance with the INUS theory of causation.

However, non-sufficiency, which underlies all remainders in conservative solutions and difficult counterfactuals in intermediate solutions, is a much stronger statement than sufficiency. While a true sufficiency statement can be based on either the truth of *A* in conjunction with the truth of *B* or the falsity of *A* alone, non-sufficiency *requires* the existence of *A* in conjunction with the absence of *B*. In consequence, both conservative and intermediate solutions regularly infer beyond the given data, whereby researchers are put at high risk of committing causal fallacies even in circumstances that are otherwise ideal for causal discovery (Baumgartner & Thiem 2017a; the proposition $A \rightarrow B$ is true for $\{A = 0, B = 1\}$, $\{A = 0, B = 0\}$, and $\{A = 1, B = 1\}$, but it is false for $\{A = 1, B = 0\}$, whereby the negation of $A \rightarrow B$, $\neg(A \rightarrow B)$, is true only for $\{A = 1, B = 0\}$). The study by Chatterley et al. (2013), who used intermediate solutions, is a case in point. As we have seen that necessary conditions cannot be causally interpreted without having demonstrated their causal relevance, the correct models at this stage of the discussion must be pruned back to models m^*_{wm} in Expression (4) and m^*_{pm} in Expression (5):

$$m^*_{wm}: \overbrace{LC \wedge NV \wedge FT}^1 \vee \overbrace{QC \wedge NV \wedge CS}^2 \vee \overbrace{QC \wedge LC \wedge CS}^3 \vee \overbrace{QC \wedge LC \wedge NV}^4 \vee \overbrace{QC \wedge LC \wedge FT}^5 \leftrightarrow WM \tag{4}$$

$$m^*_{pm}: \overbrace{li \wedge lc}^1 \vee \overbrace{cs \wedge nv \wedge li}^2 \vee \overbrace{cs \wedge nv \wedge ft \wedge lc}^3 \leftrightarrow PM \tag{5}$$

Let us now examine whether the fifth path of m^*_{wm} , the conjunction of quality construction, the presence of a local champion and familiar technology, is really a minimally sufficient condition for well-maintained school sanitation. As we can see in Table 1, this combination, which occurs for cases 6, 7, and 8, is consistently associated with the presence of the outcome. There is only one problem: the combination of quality construction and the presence of a local champion alone is also consistently associated with well-maintained school sanitation, and so is the combination of quality construction and familiar technology. But we can say even more: there is no case of quality construction that is associated with poor maintenance. Quality construction alone is consistently associated with good maintenance, making it a minimally sufficient condition for the outcome by itself. But if quality construction

itself is a minimally sufficient condition for good maintenance, then all other conditions that are part of paths 2, 3, 4, and 5 in model m^*_{wm} are redundant and, therefore, cannot possibly be demonstrated to be INUS conditions of good maintenance.

Remember that a condition, in order to be output as causally relevant under the INUS theory of causation, must be non-redundant within a sufficient condition. If a condition can be dropped from a sufficient conjunction without the latter losing its sufficiency for the outcome, causal relevance cannot be attributed to the dispensable condition. To declare something as causally relevant that cannot be shown to be causally relevant based on the empirical data thus requires the addition of artificial cases to the set of empirical cases, and this is exactly what both the conservative and the intermediate solution type do through the back door.

For example, to conclude that familiar technology was an INUS condition with respect to path 5, there would have to be a case identical with case 6 yet without the presence of both familiar technology and good maintenance, or a case identical with case 7 yet without the presence of both familiar technology and good maintenance, or a case identical with case 8 yet without the presence of both familiar technology and good maintenance. As can be verified by mere visual inspection of Table 1, none of these cases exist. In summary, neither conservative nor intermediate solutions should be used for causal data analysis, as their corresponding models assign some conditions the status of an INUS condition, although no empirical evidence for doing so exists. Only parsimonious solutions do not output such unwarranted inferences.

Eliminating alternative models from consideration

Given the corrections of the preceding section, models m^*_{wm} and m^*_{pm} must be pruned back further to yield models m'_{wm} and m'_{pm} in Expressions (6) and (7), respectively:

$$m'_{wm}: \overbrace{LC \wedge NV}^1 \vee \overbrace{QC}^2 \leftrightarrow WM \quad (6)$$

$$m'_{pm}: \overbrace{li}^1 \vee \overbrace{lc \wedge nv}^2 \leftrightarrow PM \quad (7)$$

This, of course, does not mean that the causal structures explaining the two outcomes under investigation become simpler and simpler. It only means that the given data do not permit the output of more inferences about the underlying causal structure. Implicitly, the same placeholders for unknown INUS conditions and packages of INUS conditions used above in the introductory example on the causal structure explaining *HUE* (X_1 , X_2 , and Y) remain part of m'_{wm} and m'_{pm} .

There is a third problem besetting the analysis of Chatterley and her co-authors, and this third problem has a much closer relation to QMC's roots in electrical engineering than the other two problems addressed above. More specifically, this problem consists in the import of QMC into social-scientific, causally oriented data analysis, a purpose for which QMC was never built in the first place (Thiem 2014a, 2014b; Baumgartner & Thiem 2017b). Instead, QMC was built to minimize the costs of electrical switching circuits, a purpose requiring algorithmic components that make researchers interested in causal data analysis believe that the 'story' behind their data is clean and clear, when it is in fact not. This article is not the place to go into the technical details of this problem, which have been fully laid out in Baumgartner & Thiem (2017b), and which interested readers are referred to.

For our empirical example, it means that, by using the fs/QCA software, which implements QMC with its cost-minimization objective still in operation, Chatterley *et al.* (2013) may have been presented with only the shortest of all causally interpretable models that fit their data equally well. Yet, the shortest story may not always be the correct story. When all equally well-fitting minimally necessary disjunctions of minimally sufficient conjunctions are derived from the data presented in Table 1, there are, in point of fact, 14 suitable models with respect to good maintenance and 12 such models with respect to poor maintenance, all of which are presented in Table 2, including inclusion/consistency (Incl.), raw coverage (Cov.r), and unique coverage (Cov.u) scores. Among the 14 models for good maintenance, m'_{wm} is but one option (m_8), and among the 12 models for poor maintenance, m'_{pm} is but one option (m_2) describing (in parts) the unknown causal structures Chatterley and her co-authors sought to get closer to with their study.

Again, what does this imply for policy? Imagine that the true causal structure behind good maintenance was m_{12} : the

Table 2 | QCA solutions for well-maintained school sanitation and poorly maintained school sanitation

#	WM	Incl.	Cov.r	Cov.u	m ₁	m ₂	m ₃	m ₄	m ₅	m ₆	m ₇	m ₈	m ₉	m ₁₀	m ₁₁	m ₁₂	m ₁₃	m ₁₄	
1	QC	1.00	0.875	0.00			0.375	0.125	0.500		0.375	0.500	0.500						
2	CS ^ ft	1.00	0.375	0.00						0.125							0.125	0.250	
3	CS ^ LC	1.00	0.500	0.00									0.125			0.125			
4	LC ^ NV	1.00	0.500	0.00								0.125			0.125			0.125	
5	LI ^ CS	1.00	0.625	0.00		0.125					0.125			0.125	0.250				
6	LI ^ FT	1.00	0.500	0.00					0.125					0.125	0.125	0.125	0.250	0.250	
7	LI ^ LC	1.00	0.875	0.00	0.375	0.375		0.125		0.625									
8	LI ^ NV	1.00	0.625	0.00	0.125		0.125							0.125		0.250	0.125		
PM																			
1	li	1.00	0.857	0.00		0.571			0.286	0.714	0.571	0.286	0.286	0.571					
2	cs ^ lc	1.00	0.429	0.00										0.143		0.143			
3	cs ^ qc	1.00	0.714	0.00			0.286	0.571					0.143		0.143				
4	lc ^ FT	1.00	0.429	0.00				0.286								0.143			
5	lc ^ nv	1.00	0.429	0.00		0.143													
6	lc ^ qc	1.00	0.714	0.00	0.286		0.286		0.143										
7	nv ^ qc	1.00	0.714	0.00	0.286							0.143			0.143	0.286			
8	qc ^ ft	1.00	0.429	0.00							0.143				0.143				
9	cs ^ nv ^ ft	1.00	0.286	0.00						0.143									

LI, local involvement; CS, community support; LC, local champion; NV, no vandalism; QC, quality construction; FT, familiar technology; WM, well-maintained school sanitation; PM, poorly maintained school sanitation. Lower-case letters indicate negation. Incl., inclusion; Cov.r, raw coverage; Cov.u, unique coverage.

disjunction of (1) community support and a local champion, (2) local involvement and familiar technology, and (3) local involvement and no vandalism. However, Chatterley *et al.* (2013) recommended that QC was the most important single factor. After all, QC is associated with seven out of all eight cases that showed good maintenance (raw coverage of 87.5%), and it uniquely covers more than three times as many cases as the other path of m'_{wm} , namely, the presence of a local champion and the absence of vandalism (unique coverage of 50% as against 12.5%). In that case, much money may be spent on high-end infrastructure when there was, in fact, no empirical evidence to support such a course of action, money that could have been put to much better use, for example, on incentivizing people to become local champions or on installing measures to prevent vandalism.

CONCLUSIONS

In a review article recently published in this journal, Kaminsky & Jordan (2017) have argued for the immense potential of QCA. Although the method is well positioned for building empirically founded theories that emphasize causal complexity, only very few researchers in development and planning have so far exploited the advantages QCA offers.

While this article has fully supported Kaminsky & Jordan's (2017) general call to realize the untapped potential of QCA for reaching policy-relevant insights into development and planning problems, its main purpose was to alert researchers to some consequential misapplications of this method, which have led to results in the past that misrepresented the empirical evidence marshaled to support them. By re-analyzing a recent empirical study on school sanitation maintenance in Belize, the goal of this method workshop article was thus to demonstrate how the use of QCA can be improved, which should in turn lead to more solid, evidence-based policy recommendations for development interventions.

It has been shown that necessary conditions must also prove their causal relevance, that both conservative and intermediate solutions infer beyond the data and thus create a high risk of committing causal fallacies, and that alternative, equally well-fitting models must not be eliminated before being presented to the analyst, who can then

assess or further evaluate the plausibility of these models. QCA possesses great potential for WaSH research indeed, but only correct use may systematically lead to useful policy recommendations and, ultimately, effective development interventions.

REFERENCES

- Baumgartner, M. 2008 *Regularity theories reassessed*. *Philosophia* 36 (3), 327–354.
- Baumgartner, M. & Thiem, A. 2017a *Often trusted but never (properly) tested: evaluating qualitative comparative analysis*. *Sociological Methods & Research*. Advance online publication. DOI: 10.1177/0049124117701487.
- Baumgartner, M. & Thiem, A. 2017b *Model ambiguities in configurational comparative research*. *Sociological Methods & Research* 46 (4), 954–987.
- Beebe, H., Hitchcock, C. & Menzies, P. 2009 *The Oxford Handbook of Causation*. Oxford University Press, Oxford.
- Bradshaw, Y. W., Kaiser, P. J. & Ndegwa, S. N. 1995 *Rethinking theoretical and methodological approaches to the study of African development*. *African Studies Review* 38 (2), 39–65.
- Cartwright, N. 2007 *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge University Press, Cambridge.
- Chatterley, C., Linden, K. G. & Javernick-Will, A. N. 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. N., 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 (6), 1–14.
- Drass, K. A. & Spencer, J. W. 1987 *Accounting for pre-sentencing recommendations: typologies and probation officers' theory of office*. *Social Problems* 34 (3), 277–293.
- Edwards, F. H. 1973 *The Principles of Switching Circuits*. MIT Press, Cambridge, MA.
- Gasparro, K. E. & Walters, J. P. 2017 *Revealing causal pathways to sustainable water service delivery using fsQCA*. *Journal of Water, Sanitation and Hygiene for Development* 7 (4), 546–556.
- Gerring, J. 2001 *Social Science Methodology: A Critical Framework*. Cambridge University Press, Cambridge.
- Graßhoff, G. & May, M. 2001 *Causal regularities*. In: *Current Issues in Causation* (W. Spohn, M. Ledwig & M. Esfeld, eds). Mentis, Münster, pp. 85–113.
- Hacker, M. E. & Kaminsky, J. A. 2017 *Cultural preferences for the methods and motivation of sanitation infrastructure development*. *Journal of Water, Sanitation and Hygiene for Development* 7 (3), 407–415.
- Holvoet, N. & Inberg, L. 2016 *Do gender targets and gender working groups contribute to more gender-sensitive budget*

- support? Evidence from 14 sub-Saharan African countries. *European Journal of Development Research* **28** (5), 875–892.
- Hoover, K. D. 1990 The logic of causal inference: econometrics and the conditional analysis of causation. *Economics and Philosophy* **6** (2), 207–234.
- Kaminsky, J. A. & Javernick-Will, A. N. 2014 The internal social sustainability of sanitation infrastructure. *Environmental Science & Technology* **48** (17), 10028–10035.
- Kaminsky, J. A. & Jordan, E. 2017 Qualitative comparative analysis for WASH research and practice. *Journal of Water, Sanitation and Hygiene for Development* **7** (2), 196–208.
- Lam, W. & Ostrom, E. 2010 Analyzing the dynamic complexity of development interventions: lessons from an irrigation experiment in Nepal. *Policy Sciences* **43** (1), 1–25.
- Lewin, D. & Protheroe, D. 1992 *Design of Logic Systems*, 2nd edn. Chapman & Hall, London.
- Mackie, J. L. 1965 Causes and conditions. *American Philosophical Quarterly* **2** (4), 245–264.
- Mackie, J. L. 1974 *The Cement of the Universe: A Study of Causation*. Oxford University Press, Oxford.
- Mahoney, J., Kimball, E. & Koivu, K. L. 2009 The logic of historical explanation in the social sciences. *Comparative Political Studies* **42** (1), 114–146.
- McCluskey, E. J. 1956 Minimization of Boolean functions. *Bell Systems Technical Journal* **35** (6), 1417–1444.
- McCluskey, E. J. 1965 *Introduction to the Theory of Switching Circuits*. Princeton University Press, Princeton, NJ.
- Meehl, P. E. 2001 Comorbidity and taxometrics. *Clinical Psychology: Science and Practice* **8** (4), 507–519.
- Person, M. T., Delea, M. G., Garn, J. V., Alexander, K., Abaire, B. & Freeman, M. C. 2017 Assessing the influence of social capital on water point sustainability in rural Ethiopia. *Journal of Water, Sanitation and Hygiene for Development* **7** (4), 611–622.
- Psillos, S. 2009 Regularity theories. In: *The Oxford Handbook of Causation* (H. Beebe, C. Hitchcock & P. Menzies, eds). Oxford University Press, Oxford, pp. 131–157.
- Ragin, C. C. 1987 *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press, Berkeley, CA.
- Ragin, C. C. 1989 The logic of the comparative method and the algebra of logic. *Journal of Quantitative Anthropology* **1** (4), 373–398.
- Ragin, C. C. 2000 *Fuzzy-Set Social Science*. University of Chicago Press, Chicago, IL.
- Ragin, C. C. 2008 *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press, Chicago, IL.
- Ragin, C. C. & Davey, S. 2014 *fs/QCA: Fuzzy-Set/Qualitative Comparative Analysis, Version 2.5 [Computer Program]*. Department of Sociology, University of California, Irvine, CA.
- Ragin, C. C. & Sonnett, J. 2005 Between complexity and parsimony: Limited diversity, counterfactual cases and comparative analysis. In: *Vergleichen in der Politikwissenschaft* (S. Kropp & M. Minkenberg, eds). VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 180–197.
- Ragin, C. C. & Strand, S. I. 2008 Using qualitative comparative analysis to study causal order: comment on Caren and Panofsky (2005). *Sociological Methods & Research* **36** (4), 431–441.
- Ragin, C. C., Mayer, S. E. & Drass, K. A. 1984 Assessing discrimination: a Boolean approach. *American Sociological Review* **49** (2), 221–234.
- Reiss, J. 2009 Causation in the social sciences: evidence, inference, and purpose. *Philosophy of the Social Sciences* **39** (1), 20–40.
- Rihoux, B. & Ragin, C. C. 2009 *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE, London.
- Roth, C. H. & Kinney, L. L. 2014 *Fundamentals of Logic Design*, 7th edn. Cengage Learning, Stamford, CT.
- Sawcer, S., Franklin, R. J. M. & Ban, M. 2014 Multiple sclerosis genetics. *Lancet Neurology* **13** (7), 700–709.
- Schneider, C. Q. & Wagemann, C. 2012 *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis (QCA)*. Cambridge University Press, Cambridge.
- Thiem, A. 2014a Mill's methods, induction and case sensitivity in qualitative comparative analysis: a comment on Hug (2013). *Qualitative & Multi-Method Research* **12** (2), 19–24.
- Thiem, A. 2014b Navigating the complexities of qualitative comparative analysis: case numbers, necessity relations, and model ambiguities. *Evaluation Review* **38** (6), 487–513.
- Thiem, A. 2016 Standards of good practice and the methodology of necessary conditions in qualitative comparative analysis. *Political Analysis* **24** (4), 478–484.
- Thiem, A. 2017 Conducting configurational comparative research with qualitative comparative analysis: a hands-on tutorial for applied evaluation scholars and practitioners. *American Journal of Evaluation* **38** (3), 420–435.
- Thiem, A. 2018 *QC Apro: Advanced Functionality for Performing and Evaluating Qualitative Comparative Analysis, Version 1.1-2*. <http://www.alrik-thiem.net/software/>.
- Thiem, A. & Baumgartner, M. 2016a Back to square one: a reply to Munck, Paine, and Schneider. *Comparative Political Studies* **49** (6), 801–806.
- Thiem, A. & Baumgartner, M. 2016b Modeling causal irrelevance in evaluations of configurational comparative methods. *Sociological Methodology* **46** (1), 345–357.
- Thiem, A., Baumgartner, M. & Bol, D. 2016 Still lost in translation! A correction of three misunderstandings between configurational comparativists and regression analysts. *Comparative Political Studies* **49** (6), 742–774.
- Vaisey, S. 2009 QCA 3.0: the 'Ragin revolution' continues. *Contemporary Sociology* **38** (4), 308–312.
- Wulff, H. R. 1984 The causal basis of the current disease classification. In: *Health, Disease, and Causal Explanations in Medicine* (L. Nordenfelt & B. I. B. Lindahl, eds). Springer Netherlands, Dordrecht, pp. 169–177.

First received 16 October 2017; accepted in revised form 28 June 2018. Available online 18 July 2018