

Qualitative Comparative Analysis as a Method for Evaluating Complex Cases

An Overview of Literature and a Stepwise Guide with Empirical Application

Lasse M. Gerrits,¹ Stefan Verweij²

Abstract: The complexity of many cases, such as public projects and programs, requires evaluation methods that acknowledge this complexity in order to facilitate learning. On the one hand, the complexity of the cases derives from their uniqueness and nested nature. On the other hand, there is a need to compare cases in such a way that lessons can be transferred to other (future) cases in a coherent and non-anecdotal way. One method that is making headway as a complexity-sensitive, comparative method is Qualitative Comparative Analysis (QCA). This contribution aims to explain to what extent QCA is complexity-informed, to show how it can be deployed as such, and to identify its strengths and weaknesses as an evaluation method. We will discuss the properties of complexity, provide an overview of evaluation literature about QCA, and present a simplified stepwise guide for utilizing QCA in evaluation studies.

Keywords: Complexity, Evaluation, Qualitative Comparative Analysis, Configurational Method

Qualitative Comparative Analysis als Verfahren zur Bewertung komplexer Fälle.

Eine Literaturübersicht und ein schrittweiser Leitfaden mit empirischen Beispielen

Zusammenfassung: Die Komplexität vieler Prozesse, wie z.B. öffentlicher Projekte und Programme, erfordert Evaluationsmethoden, die diese Komplexität anerkennen, damit Lernen ermöglicht werden kann. Einerseits leitet sich die Komplexität der Fälle aus deren Einzigartigkeit und verschachteltem Charakter ab. Andererseits besteht aber die Notwendigkeit, Fälle so zu vergleichen, dass Gelerntes kohärent und nicht anekdotisch auf andere (zukünftige) Fälle übertragen werden kann. Eine Methode, die als komplexitätssensitive, vergleichende Methode Fortschritte macht, ist die Qualitative Comparative Analysis (QCA). Dieser Beitrag möchte erklären, inwieweit QCA eine komplexitätskundige Methode darstellt, und zeigen, wie sie als solche eingesetzt werden kann sowie ihre Stärken und Schwächen als Evaluationsmethode identifizieren. Wir werden die Eigenschaften von Komplexität diskutieren, einen Überblick der Evaluationsliteratur über QCA bereitstellen und eine vereinfachte, stufenweise Anleitung für die Nutzung von QCA in Evaluationsstudien geben.

Schlagwörter: Komplexität, Evaluation, Qualitative Comparative Analysis, konfigurationale Methode

1 Otto-Friedrich-Universität Bamberg

2 Erasmus University Rotterdam & Otto-Friedrich-Universität Bamberg

1. Introduction

Cases feature certain specific complex properties (cf. Byrne 1998; Byrne/Callaghan 2013; Glouberman/Zimmerman 2002) that need to be addressed when evaluating them (cf. Barnes/Matka/Sullivan 2003; Bressers/Gerrits 2013; Burton/Goodlad/Croft 2006; Patton 2011; Rogers 2008; Sanderson 2000, 2002; Stame 2004; Westthorp 2012), in particular when comparing several cases. Whilst single-case studies allow analyzing the case-specific details that matter, such details often disappear when cases are compared (cf. Bamberger 2000; Bamberger/Rao/Woolcock 2010). However, it is through comparison that case complexity can be highlighted. Therefore, there is a need for methods that retain the complex details of particular cases and that compare those cases in a systematic and transparent manner. We propose Qualitative Comparative Analysis (QCA) as a complexity-informed evaluation method that addresses this need. This contribution aims to explain why QCA is a suitable method, how it can be deployed in such a way that it does justice to the complexity of the cases, and to identify its strengths and weaknesses as an evaluation method. Along the way, we present many references to literature that could be useful for readers who would like to get a deeper understanding of case complexity and QCA.

We will first discuss the properties of case complexity (Section 2). We will then introduce QCA and review its appearance in evaluation literature (Section 3). Next, we will proceed with a step-wise guide to using QCA in evaluation and discuss the ways in which these steps address the properties of complexity as identified (Section 4). We will conclude this article by identifying the strengths and weaknesses of QCA based on empirical application and review of literature (Section 5).

2. Complexity and Cases

We will first need to discuss the meaning of the term ‘complexity’. As already noted by Byrne (1998), Glouberman and Zimmermann (2002), Rogers (2011) and Gerrits (2012), amongst others, there is a real difference between the complicated and the complex.³ In this article, we specifically refer to complexity as the common denominator for the group of related theories, mechanisms and concepts identified in the realm of the complexity sciences (cf. Byrne 1998, 2004, 2005, 2011a; Byrne/Callaghan 2013; Cilliers 1998; Gell-Mann 1995; Kiel 1989; Mathews/White/Long 1999; Mitchell 2009), which has been identified as a promising avenue for evaluation research (cf. e.g. Barnes/Matka/Sullivan 2003; Byrne 2013; Forss/Marra/Schwartz 2011; Mowles 2014; Patton 2011; Rogers 2008; Sanderson 2000; Stern et al. 2012; Stern 2014; Walton 2014; Wolf-Branigin 2013). While there is no universally established list of the characteristics that belong to the complexity sciences

3 We take Mowles’ (2014) criticism regarding this distinction to heart. Nevertheless, we think that a principal distinction between simplistic and complex cases is very useful in assigning meaning to the word ‘complex’. See also Rescher (1998) for an extended discussion.

– and while we acknowledge that there may be contradictory concepts and theories within that particular realm (cf. e.g. Medd 2001; Page 2008) – there are a number of common core characteristics that resurface in most writings on complexity and social reality.⁴ From that literature, we synthesized the following dimensions that are relevant to evaluation research.

Emergence. Emergence is considered the key property of complexity. It refers to the fact that local interactions build complex patterns over time; or in other words, that complexity arises from local interactions bounded in time and place (cf. Goldstein 1999; Holland 1995, 2006; Schelling 1978; Smith/Stevens 1996). We fully agree with Elder-Vass (2005) that emergence in *social* reality is essentially synchronic. It implies that – save for simulations – there is no perceptible difference between an actual starting point and the final result, i.e. diachronic emergence. Complexity can be traced back to such local interactions, albeit in one instance.

Structure. Synchronic emergence is intimately related to structure. Structure refers to the set of relationships between the constituent elements (cf. Marion 1999; Mitchell 2009; Rescher 1998). Structure is formed through the interaction between these elements (cf. Westhorp 2012). The common view is that these relationships must have some durability in order to constitute structure (cf. Meadows 2008). Without durability, there is no structure but just a one-off collection or heap of elements (cf. Laszlo 1972). However, this does not mean that the structure as a whole is homeostatic as was originally postulated (cf. e.g. Parsons 1951). It is now acknowledged that this structure evolves over time in response to incentives, such as (policy) interventions. From this, it follows that social reality is *emergently structured* (cf. Reed/Harvey 1992) and that it consists of systemic wholes (cf. Byrne 2013) whose integrities are violated when taken apart in discrete elements (cf. Flood 1999; Gerrits/Verweij 2013). Following Byrne (2005, 2013) and Harvey (2009), we understand that real world cases operate as systemic wholes that should be studied as such. Essentially, this is an anti-reductionist argument (cf. Sanderson 2002; Sibeon 1999).

Context. The systemic wholes discussed here do not operate in a vacuum: they are in many ways related to their environment with which they interact (cf. Cilliers 2001). This interaction results in changes in the systemic whole, i.e. environmental influences, such as interventions, drive the process of emergence and become part of the emergent structure (cf. Gerrits/Verweij 2013; Harvey 2009). As such, the importance of context in understanding complexity cannot be overestimated in evaluation research (cf. Westhorp 2012). Here, we touch upon the realist argument that context determines which relationships are actualized, thus driving the direction of emergence, and the resultant structures. In other words, certain future states will only be possible under certain conditions (cf. Byrne/Callaghan 2013; Byrne 2013). This means that context is explanatory for how cases as systemic wholes emerge and develop in certain (convergent or divergent) directions over time, even if they appear similar. This is a trace of the time-asymmetric and indeterministic nature of social reality (cf. Byrne 1998; Prigogine 1997). Context determines the direction of

4 This implies that we do not concern ourselves with complexity as defined in the natural sciences.

the case and it is explanatory for why certain similar conditions bring forth different outcomes, or why dissimilar conditions can bring forth similar outcomes, i.e. multifinality and equifinality. It also implies non-linearity (cf. Kiel 1989; Kiel/Elliott 1996), a disproportional relationship between local interactions and their aggregation in the systemic whole.

Human Agency. There is the question of drawing boundaries between systemic wholes and their environment. As many authors have noted (cf. e.g. Cabrera/Colosi 2008; Cabrera/Colosi/Lobdell 2008; Cilliers 2005; Richardson/Lissack 2001), ambiguity about those boundaries is commonplace. This is why we follow Flood (1999) and talk of ‘systemic wholes’ instead of ‘systems’ and appreciate that human agency has a key role in defining those systemic wholes (cf. Byrne 2009, 2013), where it is understood that, if actors in cases operate according to their own system definitions, these definitions in fact *constitute* systemic boundaries (cf. Midgley/Munlo/Brown 1998). Indeed, the process of casing – the construction of the case object – is in itself an important step in this type of research (cf. Harvey 2009; Ragin/Becker 1992).

In conclusion, complexity is not substantive in itself (cf. Westhorp 2012). Rather, complexity is driven by complex causality, the kind of causality that is conditional (i.e. contextual) and therefore local in place and temporal in time (cf. Byrne 2009; Gerrits/Verweij 2013). Consequently, actualized causal mechanisms can differ to greater or lesser extent from case to case. Therefore, a method that highlights these context-dependent mechanisms in evaluation research is needed. One such method is Qualitative Comparative Analysis or QCA. In the next section, we will discuss how QCA and evaluation are related.

3. QCA and Evaluation

QCA, originally developed by Charles Ragin (1987, 2000, 2008), has been proposed as a complexity-informed research method that mediates between the in-depth understanding of complex cases and knowledge of context-bound generality (cf. Ragin/Shulman/Weinberg/Gran 2003). QCA has been rapidly gaining popularity (cf. Rihoux/Rezsöghazy/Bol 2011; Rihoux/Álamos-Concha/Bol/Marx/Rezsöghazy 2013). Some authors have advocated QCA as a useful evaluation method for complex cases (cf. e.g. Befani/Ledermann/Sager 2007; Befani 2013; Blackman/Wistow/Byrne 2013; Byrne 2013). In our view, QCA “[...] allows researchers to explore a particular case, to test those explanations against a larger (but limited) set of cases, and to use the results of such a test to improve our understanding of singular cases” (Gerrits, 2012: 175).

To gain an understanding of how QCA has been used in evaluation research – and how it relates to complexity as described above – we conducted a search in the bibliographical database Scopus. We used Scopus’ Journal Analyzer to select the relevant evaluation journals, simply using the query *evaluation*. The resultant selection of nineteen evaluation journals were searched with the queries ALL(*qca) or

ALL (“qualitative comparative analysis”) to ensure that all articles that, in whatever variation, mention QCA were found. This resulted in twenty-six articles in total, of which nineteen actually mentioned QCA.⁵ These are listed in Table 1.

Table 1: Overview of the 19 publications discussed in this article

<i>Authors</i>	<i>Year</i>	<i>Journal</i>	<i>Focus on QCA</i>	<i>Empirical</i>	<i>No. of Cases</i>	<i>Complexity-Informed</i>
Balthasar, A.	2009	Evaluation Review	No	No	N/A	N/A
Balthasar, A.	2006	Evaluation	Yes	Yes	10	No
Befani, B.	2013	Evaluation	Yes	Yes	28 (8 / 20)*	Yes
Befani, B., Ledermann, S., Sager, F.	2007	Evaluation	Yes	Yes	15	Yes
Blackman, T., Wistow, J., Byrne, D.	2013	Evaluation	Yes	Yes	27	Yes
Byrne, D.	2013	Evaluation	Yes	No	N/A	Yes
Costa, A.F.D., Pegado, E., Ávila, P., Coelho, A.R.	2013	Evaluation and Program Planning	No	No	N/A	N/A
Holma, K., Kontinen, T.	2011	Evaluation	No	No	N/A	N/A
Holvoet, N., Dewachter, S.	2013	American Journal of Evaluation	Yes	Yes	30	No
Jackson, S.F., Kolla, G.	2012	American Journal of Evaluation	No	No	N/A	N/A
Ledermann, S.	2012	American Journal of Evaluation	Yes	Yes	11	No
Pattyn, V., Brans, M.	2013	Canadian Journal of Program Evaluation	No	No	N/A	N/A
Sager, F., Andereggen, C.	2012	American Journal of Evaluation	Yes	Yes	17	Yes
Taut, S.	2007	Canadian Journal of Program Evaluation	No	No	N/A	N/A
Ton, G.	2012	Evaluation	No	No	N/A	N/A
Verweij, S., Gerrits, L.M.	2013	Evaluation	Yes	No	N/A	Yes
Walton, M.	2014	Evaluation and Program Planning	No	No	N/A	N/A
Williams, A.M.	2010	Evaluation and Program Planning	No	No	N/A	N/A
Yin, R.K.	2013	Evaluation	No	No	N/A	N/A

*) Concerns two different evaluation studies.

5 Other articles mentioning QCA concerned different abbreviations such as the Qualification and Curriculum Authority.

Of the nineteen articles, seven included an empirical application (cf. Balthasar 2006; Befani/Ledermann/Sager 2007; Befani 2013; Blackman/Wistow/Byrne 2013; Holvoet/Dewachter 2013; Ledermann 2012; Sager/Anderegg 2012) and two concerned solely conceptual pieces (cf. Byrne 2013; Verweij/Gerrits 2013). The other articles made references to QCA, recognizing the relevance of QCA for evaluation research, but did not pursue any detailed discussion of the approach (cf. e.g. Da Costa/Pegado/Ávila/Coelho 2013; Jackson/Kolla 2012; Pattyn/Brans 2013; Ton 2012; Walton 2014; Yin 2013). This observation indicates that QCA has found its way into the evaluation literature only recently, but rapidly; the clear majority of articles (thirteen out of nineteen) was published from 2012 onwards. We cannot claim this overview to be exhaustive, *inter alia* because it is likely that QCA evaluations have been published in non-evaluation journals (cf. e.g. Dyer 2011) and because relevant book chapters (cf. e.g. Befani/Sager 2006; Verweij/Gerrits 2012) and non-English articles (e.g. Marx 2005) were excluded in Scopus. The articles selected in our overview, though, can be considered core references about QCA and evaluation specifically. We use these references in the remainder of this section. Conducting a thorough systematic review of the literature about QCA and evaluation might become worthwhile in future years, when more evaluation studies using QCA will have been published.

QCA is an umbrella term for a variety of types, including crisp set QCA (csQCA) where conditions have binary values (cf. e.g. Befani/Ledermann/Sager 2007; Befani 2013; Blackman/Wistow/Byrne 2013; Holvoet/Dewachter 2013; Ledermann 2012), multi-value QCA (mvQCA) where conditions have discrete values (cf. e.g. Balthasar 2006; Sager/Anderegg 2012) and fuzzy-set QCA (fsQCA) where conditions can have any value between 0 and 1. It has been used in evaluation literature to comparatively study any number of cases, ranging from eight (cf. Befani 2013) up to thirty (cf. Holvoet/Dewachter 2013), aiming at the identification of causal patterns that explain the occurrence of specified outcomes in cases. These causal patterns are configurations of conditions – i.e. interacting elements – that are expected to contribute to the occurrence (or non-occurrence) of the outcome of interest, e.g. to determine the effects of a certain intervention. Evaluations with QCA have included as few as three conditions (cf. Befani 2013) up to as many as nine (cf. Blackman/Wistow/Byrne 2013; Sager/Anderegg 2012). It allows evaluators to systematically compare cases so as to identify the necessary and/or sufficient (combinations of) conditions that produce outcomes. The focus on the interaction between conditions that (do not) produce the outcome, exemplifies the generative nature of QCA (cf. Befani/Ledermann/Sager 2007) with regards to nested, emergent systemic wholes. In particular QCA's abilities to systematize and make transparent comparative evaluations, to study how conditions (e.g. interventions, context characteristics) interact with each other, and its potential for limited middle-range generalization, are appreciated by evaluation scholars (cf. Balthasar 2006; Befani/Ledermann/Sager 2007; Befani 2013; Blackman/Wistow/Byrne 2013; Holvoet/Dewachter 2013; Sager/Anderegg 2012). These abilities coincide with the complexity-informed understanding of the emergently-structured nature of social reality.

Although most publications in evaluation journals recognize QCA's sensitivity to complexity – most prominently: QCA offers a more holistic, important and valuable alternative to linear approaches that try to isolate the effect of conditions (e.g. interventions or contextual characteristics) on outcomes – fewer authors have explicitly discussed and presented QCA as an approach to studying complex systemic wholes (cf. Blackman/Wistow/Byrne 2013; Byrne 2013; Verweij/Gerrits 2013).

Byrne and colleagues, for instance, building on their previous work (cf. e.g. Byrne 2005, 2009, 2011b), have conceptualized QCA as a way to study cases as complex systems, consisting of elements that interact with the system's context and producing system-level outcomes, where this level is understood in terms of emergent complexity as described in the previous section. The system elements and context characteristics are conceptualized as conditions. As such, QCA allows evaluators to trace the complex causal links between the elements and characteristics so as to explain the occurrence of outcomes. For example, Blackman, Wistow and Byrne (2013) have examined how interventions in so-called 'spearhead areas' in England worked in different contexts to (not) produce a narrowing gap of teenage conceptions relative to the country's average.

Verweij and Gerrits (2013), likewise Byrne and colleagues, commence with a conception of reality composed of multiple, nested and interacting complex systems. They thus assign systemic qualities to reality, and then go on formulating a number of epistemological – stemming from this complex understanding of reality – and consequent methodological requisites for complexity-informed evaluations. They assess QCA's properties against these requisites and ultimately conclude that QCA is complexity-informed, but that its essentially a-temporal nature poses challenges for using QCA to study complex systems and phenomena, something that Byrne agrees with (cf. Byrne/Callaghan 2013). We note that Verweij and Gerrits' discussion links with Realistic Evaluation (RE). As they explain in a related publication (2013), complexity is rooted in the critical realist ontology. RE is also rooted in critical realism (cf. Pawson/Tilley 1997), and the parallels between QCA and RE, in turn, have been highlighted by e.g. Befani, Ledermann and Sager (2007) and Sager and Andereggen (2012). In the next section we will build on the works discussed here, explaining and showing how QCA can be used for complexity-informed evaluation.

4. QCA for Complexity-Informed Evaluation

The selection of cases and the identification of the conditions that are included in the analysis are pivotal in any case study research in general, and in QCA in particular. The conditions that are included constitute and define the case. As can be surmised from our discussion above, what constitutes a case – i.e. which conditions are included; casing – can and should be subject to discussion.

Many select conditions on the basis of theory, for instance in order to test configurational hypotheses (cf. Amenta/Poulsen 1994), and the QCA literature general-

ly agrees that this is a good practice (cf. Schneider/Wagemann 2010). In our view, theory here includes policies as well because policies are de facto theories of how an intervention in a certain context produces an outcome, i.e. the so-called policy theory (cf. Leeuw 2003). It is also agreed in the QCA literature that the selection of conditions depends on empirical knowledge of the cases (cf. Schneider/Wagemann 2010). In our view, it is essential to develop the cases from explorative observations and discussions in the field because this will guide the researcher towards the local interactions and boundary judgments that build complex systemic wholes. Rantala and Hellström (2001) and Verweij and Gerrits (2015), for instance, developed conditions in a grounded manner from qualitative interviews that focused on the human agency of the respondents. Rantala and Hellström interviewed teenagers and from the transcripts, via coding, built categories of how the teenagers reflect upon the world and themselves through practicing art. Similarly, Verweij and Gerrits (2015) interviewed managers of a large infrastructure project, using an open approach that focused on the day-to-day actions of the managers, and through iterative coding constructed categories of management conditions that produce satisfactory outcomes in the implementation of the project. Such grounded and iterative approaches to understanding cases can be recognized in other types of evaluation as well.

It is important that cases can be expected, in principle, to produce the outcome of interest. For Verweij and Gerrits (2015) this meant that the interviewed managers provided cues about what the outcome of their management actions was. In other words, they highlighted the emergent nature of the complex projects they were working on. Of course, it is not necessary that the outcome is actually produced: theories may be wrong, policies can (partly) fail and management actions may not work. The question is under what conditions the outcomes are produced or not.

Developing the cases from explorative observations and discussions in the field is a particularly useful approach for complexity-informed QCA evaluation. Although there are authors who argue that cases can be approached quantitatively, e.g. using Configuration-Frequency-Analysis – and we acknowledge the value of that for certain research goals – we would like to stress that case-based research first and foremost involves qualitative, in-depth, data. As e.g. Checkland (1981), Uprichard and Byrne (2006) and Wagenaar (2007) have argued, we can only get to know the operation of complex systemic wholes (note the plural here) through the views, words and actions of the people working within those. They guide the researcher in generating the boundaries of the wholes and in identifying the causal relationships within those cases (cf. e.g. Byrne 2011a), as such addressing the dimensions of emergence and human agency. In fact, cases display a level of detail and complexity that are extremely hard to define by researchers without the help of respondents (cf. Wagenaar 2007). Qualitative data help preserving as much of the actual complexity of structure and emergence as possible in a coherent way. Note that this also implies that the researcher has actual in-depth knowledge of each individual case, which is important in the calibration and interpretation processes in QCA (cf. Schneider/Wagemann 2010).

Having determined the conditions and the outcome of interest and having collected the data, the next steps of the procedure are all geared towards a transparent

reduction of complexity. This reduction is required for the systematic comparison. We stress and will demonstrate below that this reduction is complexity-informed, not the least because the data collection as described above is geared towards preserving maximum complexity.

First, the researcher should decide about the type of QCA. The three available main types – csQCA, mvQCA and fsQCA, see above – all have certain advantages, depending on the aims of the research and the types of conditions. Explanations of the different types are available in literature (cf. e.g. Rihoux/Ragin 2009; Schneider/Wagemann 2012), as well as discussions about their relative advantages and disadvantages (cf. e.g. Thiem 2013; Vink/Van Vliet 2009, 2013). Ostensibly, a complexity-informed analysis would benefit from mvQCA or fsQCA because these types allow a more fine-grained, detailed description of conditions than csQCA does: the first allows multiple discrete values (e.g. 0, 1 and 2) and the latter allows any values between 0 and 1. However, these types can introduce a faux precision that is not always present in the real world. Furthermore, it is argued, for instance by Byrne (2013), that it concerns the differences in kind and not the differences in degree that policy makers intend to establish with their interventions, and which thus should be the focus of the evaluation. Moreover and specifically for fsQCA, especially when cases (hence conditions) are developed from explorative observations and discussions in the field, it seems hard to explicate the difference between fuzzy-set scores of e.g. 0.3 and 0.4. The choice for the type of QCA ultimately depends on the research question and the nature of the data at hand. Verweij and Gerrits (2015), for instance, opted for mvQCA because they reasoned that the categories that were developed from the qualitative data were best described in discrete terms.

The next step is to structure the qualitative data into a data matrix. This involves calibration: the process of actually assigning set scores to cases (cf. Schneider/Wagemann 2012). The data matrix lists all cases (first column), the scores of the cases on the conditions (middle columns), and the scores of the cases on the outcome condition (final column). The data matrix expresses the interactions between elements (e.g. interventions) and context characteristics in terms of configurations of conditions. Each row shows the particular interactions for each case. At this stage, the researcher will note the contextual and emergent nature of the nested cases, i.e. how outcomes are stronger or weaker associated with certain configurations.

Next, the data matrix needs to be converted into a so-called truth table. The truth table lists all the logically possible combinations of conditions (excluding the outcome) and sorts the cases accordingly. By definition, each case can only be assigned to one configuration. Each configuration is then scored on the outcome condition in the final column of the truth table. For illustrative purposes we included Table 2 as an example of an mvQCA truth table (adapted from Verweij/Gerrits 2015).

Table 2: Multi-Value QCA Truth Table

Event	Management	Cooperation	Satisfaction	N	Cases
0	0	2	1	1	TUN
1	0	1	0	2	BI1, TRA
1	1	0	C	5	BI2, LAN, ZON, LEE, PRO
1	1	2	C	3	WAT, RIJ, CIV
1	0	0	1	1	ENV
1	1	1	1	1	WES
0	0	1	0	4	CR1, SOI, BAD, THE
0	0	0	1	1	CR2

The authors of this truth table conducted a complexity-informed evaluation of the project management of a large infrastructure project, focusing on how public and private managers managed and cooperated in different events-as-cases, and which contextualized combinations of management strategies and cooperative strategies produced good outcomes (satisfaction) and which ones did not. Whereas the data matrix shows the configuration for each case, the truth table shows the logically possible configurations and the cases that represent them. In other words, the matrix and the truth table *represent* the same qualitative rich data, but whereas the first focuses on the uniqueness and nested nature of the cases, the latter compares them, stressing their similarities. Importantly, the truth table exhibits that the context determines the cases' trajectories in terms of multifinality and equifinality. That is, Table 2 shows that different combinations of conditions produce the same outcome, and that similar conditions produce different outcomes depending on the context of other conditions.

Note the 'C's' in the outcome column of the table. This signals that the same configuration produces different outcomes in cases, i.e. a contradiction. This is another feature of QCA's complexity-informed nature. It draws our attention to the possibility that the cases' boundaries were defined too narrow, i.e. that there are additional conditions that explain why the cases in the third and fourth configurations of the truth table produce contradicting outcomes. This urges the researcher to continue the process of casing. Note further that not all logically possible combinations are represented by a case. That is, only eight configurations are empirically present in Table 2, whilst the number of combinations that are logically possible is twelve (i.e. $2^2 \cdot 1^3$). This is called limited diversity. It could be considered a problem because it limits the options for the further reduction of complexity in the next step of the QCA process. Alternatively, we can understand the empirical absence of configurations as indicative for the nestedness of cases, that cases have unique properties but also exhibit similarities which makes certain configurations less likely. Byrne (2013) asserts this is an indication of path-dependency.

The final step in the reduction of complexity is the development of the minimized solution formula in order to derive the persistent patterns across cases. This formula contains the information presented in the truth table, but reduced to one

formula using Boolean operators. For the outcome [1] of the truth table, including the contradictory rows for illustrative purposes as if they also produced the outcome, this formula is shown in Table 3 (adapted from Verweij/Gerrits 2015). The * symbol indicates logical-AND and + symbol indicates logical-OR.

Table 3: Minimized Solution Formula

Generalized Pathways			OUT
			[1]
Conditions	MAN [0] * COOP [0] +	EVENT [0] * MAN [0] * COOP [2] +	EVENT [1] * MAN [1]
Cases	ENV, CR2	TUN	WES, BI2, LAN, ZON, LEE, PRO, WAT, RIJ, CIV

The minimized solution formula is arrived at through pairwise comparison of configurations that agree on the outcome and differ in but one other condition. For example, the configurations represented by the cases ENV and CR2 that were combined into “MAN [0] * COOP [0]”, allowed the researchers to conclude that the combination of this project management strategy (MAN) with this cooperation strategy (COOP) works in both contexts “Event [0]” OR “Event [1]”. The minimization procedure results in generalized pathways that simultaneously exhibit the cases’ multifinal and equifinal nature, as shown in Table 2.

It is important to stress here that the minimized solution formula is not the final step in the QCA process. We would like to emphasize that the solution formula in itself is of little value without a (qualitative) interpretation of the results. Two measures guide the interpretation: consistency and coverage (cf. Ragin 2006). The first one is a measure for assessing theoretical (subset-relational) strength and the second one for assessing empirical strength. For instance, the consistencies of the two configurations in Table 2 that exhibit contradictions are lower than the configurations that have no contradictions. Consequently, the generalized path EVENT [1] * MAN [1] has a lower consistency than the other two paths in Table 2. However, the coverage of the path is higher than that of the other two, because it covers more cases. These measures help the evaluator to assess and communicate more clearly what works, in what contexts and to what extent (cf. Rogers 2011). Apart from these measures, as the results of the QCA are ultimately rooted in the richly detailed systemic wholes, not interpreting the results in light of these wholes runs the risk of severe misinterpretation. When researchers communicate the QCA results as straightforward recipes for changing policies or management practices in order to improve them, the opposite could actually be accomplished. It goes without saying that this cannot be the purpose of evaluation.

As can be surmised from this discussion: QCA is a decidedly iterative approach where the evaluator is allowed to move between the steps, e.g. to solve limited diversity and contradictions, and where she/he is encouraged to adjust the whole reduction procedure according to the interpretation of the results. The fact that QCA

can be used to both explore and test patterns is testimony to its versatility in that respect. As such, QCA first and foremost represents a research *cycle* between the considerable complexity of individual cases and the simplification necessary to compare patterns across cases. As many (cf. e.g. Cilliers 2002, 2005) have noted: simplification and reductionism are inevitable when understanding complex cases. However, there is a real difference between the practical constraints in allowing observing full complexity, and reductionism as expressed in Occam's razor. QCA is a method that accounts for the real-world complexity during the comparative stage and as such aids learning through evaluation (cf. Sanderson 2000).

5. Conclusions

Complexity appears to be increasingly acknowledged in evaluation studies. This paper made the case for QCA as one of the approaches that complexity-informed evaluators can use. This was done by discussing the properties of complexity in relation to the properties of QCA, explaining how context and human agency are central to evaluation research, and showing that QCA can be utilized to articulate these characteristics when evaluating complex cases. Inevitably, the complexity of the individual cases is reduced in (and because of) the comparative process, but it retains traces of that complexity in terms of contexts, equifinality and multifinality, i.e. emergence, structure and trajectory. Closing the cycle through a qualitative interpretation of the solution formula assigns meaning to the pertinent traces of complexity present in that formula.

While we make the case for QCA, we do not mean to say that QCA is the single best method for analyzing and comparing complex cases. The formal procedures in QCA, i.e. from the data matrix up to the minimized solution formula, are a-temporal: there is no temporal order in the conditions that are included in the comparison. Although solutions to this problem have been suggested in literature (cf. Caren/Panofsky 2005; Hak/Jaspers/Dul 2013; Hino 2009; Ragin/Strand 2008), it is still the case that QCA in itself is not really appropriate for studying temporal dynamics in cases (cf. Verweij/Gerrits 2013). Narrative approaches, longitudinal approaches, such as path-dependency and even-sequence analyses, for instance, may be more suitable here. The synchronic nature of social emergence, as argued here, highlights the importance of addressing human agency in the conduction of evaluation with QCA because it is through the eyes of people in cases that the case's boundaries and emergent properties can be reconstructed. The approach is not a mechanical exercise and including the evaluatees in the application is key to understanding the solution formula. As such, QCA is not just a tool but rather an evaluation logic that helps the researcher focusing on the complexity of the case.

6. References

- Amenta, Edwin/Poulsen, Jane D. (1994): Where to Begin: A Survey of Five Approaches to Selecting Independent Variables for Qualitative Comparative Analysis. In: *Sociological Methods and Research*, 23 (1), pp. 22-53.
- Balthasar, Andreas (2006): The Effects of Institutional Design on the Utilization of Evaluation: Evidence Using Qualitative Comparative Analysis (QCA). In: *Evaluation*, 12 (3), pp. 353-371.
- Bamberger, Michael (Ed.) (2000): *Directions in Development: Integrating Quantitative and Qualitative Research in Development Projects*. Washington: The World Bank.
- Bamberger, Michael/Rao, Vijayendra/Woolcock, Michael (2010): *Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development*. Washington: The World Bank.
- Barnes, Marian/Matka, Elizabeth/Sullivan, Helen (2003): Evidence, Understanding and Complexity: Evaluation in Non-Linear Systems. In: *Evaluation*, 9 (3), pp. 265-284.
- Befani, Barbara (2013): Between Complexity and Generalization: Addressing Evaluation Challenges with QCA. In: *Evaluation*, 19 (3), pp. 269-283.
- Befani, Barbara/Ledermann, Simone/Sager, Fritz (2007): Realistic Evaluation and QCA: Conceptual Parallels and an Empirical Application. In: *Evaluation*, 13 (2), pp. 171-192.
- Befani, Barbara/Sager, Fritz (2006): QCA as a Tool for Realistic Evaluations: The Case of the Swiss Environmental Impact Assessment. In: Rihoux, Benoit/Grimm, Heike (Eds.): *Innovative Comparative Methods for Policy Analysis: Beyond the Quantitative-Qualitative Divide*. New York: Springer, pp. 263-284.
- Blackman, Tim/Wistow, Jonathan/Byrne, Dave S. (2013): Using Qualitative Comparative Analysis to Understand Complex Policy Problems. In: *Evaluation*, 19 (2), pp. 126-140.
- Bressers, Nanny E. W./Gerrits, Lasse M. (2013): A Complexity-Informed Approach to Evaluating National Knowledge and Innovation Programmes. In: *Systems Research and Behavioral Science*, 32 (1), pp. 50-63.
- Burton, Paul/Goodlad, Robina/Croft, Jacqui (2006): How Would We Know What Works? Context and Complexity in the Evaluation of Community Involvement. In: *Evaluation*, 12 (3), pp. 294-312.
- Byrne, David S. (1998): *Complexity Theory and the Social Sciences: An Introduction*. New York: Routledge.
- Byrne, David S. (2004): *Complexity Theory and Social Research*. In: *Social Research Update*. Verfügbar unter: <http://sru.soc.surrey.ac.uk/SRU18.html> [17.02.2016].
- Byrne, David S. (2005): Complexity, Configurations and Cases. In: *Theory, Culture & Society*, 22 (5), pp. 95-111.
- Byrne, David S. (2009): Complex Realist and Configurational Approaches to Cases: A Radical Synthesis. In: David S. Byrne/Charles C. Ragin (Eds.): *The Sage Handbook of Case-Based Methods*. London: Sage, pp. 101-119.
- Byrne, David S. (2011a): *Applying Social Science: The Role of Social Research in Politics, Policy and Practice*. Bristol: The Policy Press.
- Byrne, David S. (2011b): Exploring Organizational Effectiveness: The Value of Complex Realism as a Frame of Reference and Systematic Comparison as a Method. In: Allen, Peter/Maguire, Steve/McKelvey, Bill (Eds.): *The Sage Handbook of Complexity and Management*. London: Sage, pp. 131-141.
- Byrne, David S. (2013): Evaluating Complex Social Interventions in a Complex World. In: *Evaluation*, 19 (3), pp. 217-228.
- Byrne, David S./Callaghan, Gillian (2013): *Complexity Theory and the Social Sciences: The State of the Art*. Abingdon: Routledge.
- Cabrera, Derek/Colosi, Laura (2008): Distinctions, Systems, Relationships, and Perspectives (DSRP): A Theory of Thinking and of Things. In: *Evaluation and Program Planning*, 31 (3), pp. 311-317.
- Cabrera, Derek/Colosi, Laura/Lobdell, Claire (2008): Systems Thinking. In: *Evaluation and Program Planning*, 31 (3), pp. 299-310.
- Caren, Neal/Panofsky, Aaron (2005): TQCA: A Technique for Adding Temporality to Qualitative Comparative Analysis. In: *Sociological Methods and Research*, 34 (2), pp. 147-172.
- Checkland, Peter (1981): *Systems Thinking, Systems Practice*. Chichester: John Wiley & Sons.
- Cilliers, Paul (1998): *Complexity and Postmodernism: Understanding Complex Systems*. New York: Routledge.

- Cilliers, Paul (2001): Boundaries, Hierarchies and Networks in Complex Systems. In: *International Journal of Innovation Management*, 5 (2), pp. 135-147.
- Cilliers, Paul (2002): Why We Cannot Know Complex Things Completely. In: *Emergence*, 4 (1-2), pp. 77-84.
- Cilliers, Paul (2005): Knowledge, Limits and Boundaries. In: *Futures*, 37 (7), pp. 605-613.
- Da Costa, António Firminio/Pegado, Elsa/Ávila, Patrícia/Coelho, Ana R. (2013): Mixed-Methods Evaluation in Complex Programmes: The National Reading Plan in Portugal. In: *Evaluation and Program Planning*, 39, pp. 1-9.
- Dyer, Wendy (2011): Mapping Pathways. In: Williams, Malcolm/Vogt, Paul (Eds.): *The SAGE Handbook of Innovation in Social Research Methods*. New York: Sage Publications.
- Elder-Vass, Dave (2005): Emergence and the Realist Account of Cause. In: *Journal of Critical Realism*, 4 (2), pp. 315-338.
- Flood, Robert L. (1999): *Rethinking the Fifth Discipline: Learning within the Unknowable*. London: Routledge.
- Forss, Kim/Marra, Mita/Schwartz, Robert (Eds.) (2011): *Evaluating the Complex: Attribution, Contribution, and beyond*. New Brunswick: Transaction Publishers.
- Gell-Mann, Murray (1995): What is Complexity? Remarks on Simplicity and Complexity by the Nobel Prize-Winning Author of the Quark and the Jaguar. In: *Complexity*, 1 (1), pp. 16-19.
- Gerrits, Lasse M. (2012): *Punching Clouds: An Introduction to the Complexity of Public Decision-Making*. Litchfield Park: Emergent Publications.
- Gerrits, Lasse M./Verweij, Stefan (2013): Critical Realism as a Meta-Framework for Understanding the Relationships Between Complexity and Qualitative Comparative Analysis. In: *Journal of Critical Realism*, 12 (2), pp. 166-182.
- Glouberman, Sholom/Zimmerman, Brenda (2002): *Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?*. Toronto: Commission on the Future of Health Care in Canada.
- Goldstein, Jeffrey (1999): Emergence as a Construct: History and Issues. In: *Emergence*, 1 (1), pp. 49-72.
- Hak, Tony/Jaspers, Ferdinand/Dul, Jan (2013): The Analysis of Temporally Ordered Configurations: Challenges and Solutions. In: Fiss, Peer C./Cambré, Bart/Marx, Axel (Eds.): *Configurational Theory and Methods in Organizational Research*. Bingley: Emerald, pp. 107-127.
- Harvey, David L. (2009): Complexity and Case. In: Byrne, David S./Ragin, Charles C. (Eds.): *The Sage Handbook of Case-Based Methods*. London: Sage, pp. 15-38.
- Hino, Airo (2009): Time-Series QCA: Studying Temporal Change through Boolean Analysis. In: *Sociological Theory and Methods*, 24 (2), pp. 247-265.
- Holland, John H. (1995): *Hidden Order: How Adaptation Builds Complexity*. New York: Basic Books.
- Holland, John H. (2006): Studying Complex Adaptive Systems. In: *Journal of Systems Science and Complexity*, 19 (1), pp. 1-8.
- Holvoet, Nathalie/Dewachter, Sara (2013): Multiple Paths to Effective National Evaluation Societies: Evidence from 37 Low- and Middle-Income Countries. In: *American Journal of Evaluation*, 34 (4), pp. 519-544.
- Jackson, Suzanne F./Kolla, Gillian (2012): A new Realistic Evaluation Analysis Method: Linking Coding of Context, Mechanism, and Outcome Relationships. In: *American Journal of Evaluation*, 33 (3), pp. 339-349.
- Kiel, L. Douglas (1989): Nonequilibrium Theory and its Implications for Public Administration. In: *Public Administration Review*, 49 (6), pp. 544-551.
- Kiel, L. Douglas/Elliott, Euel (Eds.) (1996): *Chaos Theory in the Social Sciences: Foundations and Applications*. Ann Arbor: The University of Michigan Press.
- Laszlo, Ervin (1972): *The Systems View of the World: The Natural Philosophy of the New Developments in the Sciences*. New York: George Braziller.
- Ledermann, Simone (2012): Exploring the Necessary Conditions for Evaluation Use in Program Change. In: *American Journal of Evaluation*, 33 (2), pp. 159-178.
- Leeuw, Frans L. (2003): Reconstructing Program Theories: Methods Available and Problems to Be Solved. In: *American Journal of Evaluation*, 24 (1), pp. 5-20.
- Marion, Russ (1999): *The Edge of Organization: Chaos and Complexity Theories of Formal Social Systems*. Thousand Oaks: Sage.

- Marx, Axel (2005): Systematisch comparatief case onderzoek van evaluatieonderzoek. In: *Tijdschrift Voor Sociologie*, 26 (1), pp. 95-113.
- Mathews, K. Micheal/White, Michael C./Long, Rebecca G. (1999): Why Study the Complexity Sciences in the Social Sciences?. In: *Human Relations*, 52 (4), pp. 439-462.
- Meadows, Donella H. (2008): *Thinking in Systems: A Primer*. White River Junction: Chelsea Green.
- Medd, Will (2001): Making (Dis)Connections: Complexity and the Policy Process. In: *Social Issues*, 1 (2).
- Midgley, Gerald/Munlo, Isaac/Brown, Mandy (1998): The Theory and Practice of Boundary Critique: Developing Housing Services for Older People. In: *The Journal of the Operational Research Society*, 49 (5), pp. 467-478.
- Mitchell, Sandra (2009): Complexity and Explanation in the Social Sciences. In: Mantzavinos, Chrysostomos (Ed.): *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*. Cambridge: Cambridge University Press, pp. 130-145.
- Mowles, Chris (2014): Complex, but Not Quite Complex Enough: The Turn to the Complexity Sciences in Evaluation Scholarship. In: *Evaluation*, 20 (2), pp. 160-175.
- Page, Scott E. (2008): Uncertainty, Difficulty, and Complexity. In: *Journal of Theoretical Politics*, 20 (2), pp. 115-149.
- Parsons, Talcott (1951): *The Social System*. New York: The Free Press.
- Patton, Michael Q. (2011): *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: The Guilford Press.
- Pattyn, Valérie/Brans, Marleen (2013): Outsource versus In-House? An Identification of Organizational Conditions Influencing the Choice for Internal or External Evaluators. In: *The Canadian Journal of Program Evaluation*, 28 (2), pp. 43-63.
- Pawson, Ray/Tilley, Nick (1997): *Realistic Evaluation*. London: Sage.
- Prigogine, Ilya (1997): *The End of Certainty: Time, Chaos, and the New Laws of Nature*. New York: The Free Press.
- Ragin, Charles C. (1987): *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Los Angeles: University of California Press.
- Ragin, Charles C. (2000): *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
- Ragin, Charles C. (2006): Set Relations in Social Research: Evaluating their Consistency and Coverage. In: *Political Analysis*, 14 (3), pp. 291-310.
- Ragin, Charles C. (2008): *Redesigning Social Inquiry: Fuzzy Sets and beyond*. Chicago: University of Chicago Press.
- Ragin, Charles C./Becker, Howard S. (Eds.) (1992): *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press.
- Ragin, Charles C./Shulman, David/Weinberg, Adam/Gran, Brian (2003): Complexity, Generality, and Qualitative Comparative Analysis. In: *Field Methods*, 15 (4), pp. 323-340.
- Ragin, Charles C./Strand, Sarah I. (2008): Using Qualitative Comparative Analysis to Study Causal Order: Comment on Caren and Panofsky (2005). In: *Sociological Methods and Research*, 36 (4), pp. 431-441.
- Rantala, Kati/Hellström, Eeva (2001): Qualitative Comparative Analysis and a Hermeneutic Approach to Interview Data. In: *International Journal of Social Research Methodology*, 4 (2), pp. 87-100.
- Reed, Michael/Harvey, David L. (1992): The New Science and the Old: Complexity and Realism in the Social Sciences. In: *Journal for the Theory of Social Behaviour*, 22 (4), pp. 353-380.
- Rescher, Nicholas (1998): *Complexity: A Philosophical Overview*. New Brunswick: Transaction Publishers.
- Richardson, Kurt A./Lissack, Michael R. (2001): On the Status of Boundaries, Both Natural and Organizational: A Complex Systems Perspective. In: *Emergence*, 3 (4), pp. 32-49.
- Rihoux, Benoît/Álamos-Concha, Priscilla/Bol, Damien/Marx, Axel/Rezsöhazy, Ilona (2013): From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011. In: *Political Research Quarterly*, 66 (1), pp. 175-184.
- Rihoux, Benoît/Ragin, Charles C. (Eds.) (2009): *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. London: Sage.
- Rihoux, Benoît/Rezsöhazy, Ilona/Bol, Damien (2011): Qualitative Comparative Analysis (QCA) in Public Policy Analysis: An Extensive Review. In: *German Policy Studies*, 7 (3), pp. 9-82.
- Rogers, Patricia J. (2008): Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. In: *Evaluation*, 14 (1), pp. 29-48.

- Rogers, Patricia J. (2011): Implications of Complicated and Complex Characteristics for Key Tasks in Evaluation. In: Forss, Kim/Marra, Mita/Schwartz, Robert (Eds.): *Evaluating the Complex: Attribution, Contribution, and Beyond*. New Brunswick: Transaction Publishers, pp. 33-52.
- Sager, Fritz/Andereggen, Céline (2012): Dealing with Complex Causality in Realist Synthesis: The Promise of Qualitative Comparative Analysis. In: *American Journal of Evaluation*, 33 (1), pp. 60-78.
- Sanderson, Ian (2000): Evaluation in Complex Policy Systems. In: *Evaluation*, 6 (4), pp. 433-454.
- Sanderson, Ian (2002): Evaluation, Policy Learning and Evidence-Based Policy Making. In: *Public Administration*, 80 (1), pp. 1-22.
- Schelling, Thomas C. (1978): *Micromotives and Macrobehavior*. New York: W.W. Norton & Company.
- Schneider, Carsten Q./Wagemann, Claudius (2010): Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy Sets. In: *Comparative Sociology*, 9 (3), pp. 397-418.
- Schneider, Claudius Q./Wagemann, Carsten (2012): *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.
- Sibeon, Roger (1999): Anti-Reductionist Sociology. In: *Sociology*, 33 (2), pp. 317-334.
- Smith, Thomas S./Stevens, Gregory T. (1996): Emergence, Self-Organisation, and Social Interaction: Arousal Dependent Structure in Social Systems. In: *Sociological Theory*, 14 (2), pp. 131-153.
- Stame, Nicoletta (2004): Theory-Based Evaluation and Types of Complexity. In: *Evaluation*, 10 (1), pp. 58-76.
- Stern, Elliot (2014): Editorial. In: *Evaluation*, 20 (2), pp. 157-159.
- Stern, Elliot et al. (2012): *Broadening the Range of Designs and Methods for Impact Evaluations: Report of a Study Commissioned by the Department for International Development*. London: Department for International Development.
- Thiem, Alrik (2013): Clearly Crisp, and Not Fuzzy: A Reassessment of the (Putative) Pitfalls of Multi-Value QCA. In: *Field Methods*, 25 (2), pp. 197-207.
- Ton, Giel (2012): The Mixing of Methods: A Three-Step Process for Improving Rigour in Impact Evaluations. In: *Evaluation*, 18 (1), pp. 5-25.
- Uprichard, Emma/Byrne, David S. (2006): Representing Complex Places: A Narrative Approach. In: *Environment and Planning A*, 38 (4), pp. 665-676.
- Verweij, Stefan/Gerrits, Lasse M. (2012): Assessing the Applicability of Qualitative Comparative Analysis for the Evaluation of Complex Projects. In: Lasse M. Gerrits/Peter K. Marks (Eds.): *Compact 1: Public Administration in Complexity*. Litchfield Park: Emergent Publications, pp. 93-117.
- Verweij, Stefan/Gerrits, Lasse M. (2013): Understanding and Researching Complexity with Qualitative Comparative Analysis: Evaluating Transportation Infrastructure Projects. In: *Evaluation*, 19 (1), pp. 40-55.
- Verweij, S./Gerrits, L.M. (2015): How Satisfaction is Achieved in the Implementation Phase of Large Transportation Infrastructure Projects: A Qualitative Comparative Analysis into the A2 Tunnel Project. In: *Public Works Management & Policy*, 20 (1), pp. 5-28.
- Vink, Maarten P./Van Vliet, Olaf (2009): Not Quite Crisp, Not yet Fuzzy? Assessing the Potentials and Pitfalls of Multi-Value QCA. In: *Field Methods*, 21 (3), pp. 265-289.
- Vink, Maarten P./Van Vliet, Olaf (2013): Potentials and Pitfalls of Multi-Value QCA: Response to Thiem. In: *Field Methods*, 25 (2), pp. 208-213.
- Wagenaar, Hendrik (2007): Governance, Complexity, and Democratic Participation: How Citizens and Public Officials Harness the Complexities of Neighborhood Decline. In: *The American Review of Public Administration*, 37 (1), pp. 17-50.
- Walton, Mat (2014): Applying Complexity Theory: A Review to Inform Evaluation Design. In: *Evaluation and Program Planning*, 45, pp. 119-126.
- Westhorp, Gill (2012): Using Complexity-Consistent Theory for Evaluating Complex Systems. In: *Evaluation*, 18 (4), pp. 405-420.
- Wolf-Branigin, Michael (2013): *Using Complexity Theory for Research and Program Evaluation*. Oxford: Oxford University Press.
- Yin, Robert K. (2013): Validity and Generalization in Future Case Study Evaluations. In: *Evaluation*, 19 (3), pp. 321-332.