Configurational Research with Qualitative Comparative Analysis (QCA) and Coincidence Analysis (CNA)

25-29 September 2017, Regenstrief Institute, Indianapolis, USA

Prof. Dr. Michael Baumgartner, University of Geneva, Switzerland
Dr. Alrik Thiem, University of Geneva, Switzerland

1 Workshop Description

This workshop offers an intensive 1-week introduction for health researchers to Qualitative Comparative Analysis (QCA) and Coincidence Analysis (CNA)—the two most prominent configurational comparative methods (CCMs) of causal discovery. Participants will be guided through the nuts and bolts of configurational comparative data analysis as well as cutting-edge methodological innovations. In replicating published studies from various areas of the health sciences, they will also learn how to make the most of current software for QCA and CNA.

From the philosophical roots of regularity theories of causation, over the procedural protocol of QCA, to the empirical analysis of causal chain structures with CNA, this seminar will go way beyond the material taught in other courses, enabling participants to perform QCA and CNA in a correct and sophisticated manner. It will be shown, for example, why the vast majority of past QCA studies have run the risk of failing to find the underlying causal model, why the so-called “conservative solution” in QCA is not conservative at all, what to do when data are beset by model ambiguities, and how to perform sensitivity tests. Last but not least, students will learn why recent high-profile critiques discouraging the use of QCA for empirical data analysis and causal inference lack traction. Individual consultation sessions will be offered in addition to help participants with the methodological aspects of their own research projects.
The two instructors are among the most active researchers in the field of configurational comparative data analysis. As authors of the most powerful software for QCA and CNA currently available, they also have an unrivalled familiarity with these tools, which will be made ample use of throughout the workshop.

2 Instructor Details

Michael Baumgartner is a Swiss National Science Foundation Professor at the Department of Philosophy of the University of Geneva, Switzerland. His research focuses on questions in the philosophy of science and logic, more specifically, on aspects of causation and causal reasoning with QCA and CNA, regularity theories, interventionism, determinism, logical formalization, argument reconstruction and modelling in the social sciences. He has published in journals such as the British Journal for the Philosophy of Science, Comparative Political Studies, Dialectica, Erkenntnis, Field Methods, Journal of Philosophical Logic, Sociological Methodology, Sociological Methods & Research and Synthese. He has developed the method of CNA and is a co-author of the corresponding cna package for the R environment. For further biographical details, see his Homepage.

Alrik Thiem is a post-doctoral research associate at the Department of Philosophy of the University of Geneva, Switzerland. The main part of his work addresses questions of methodology in the field of empirical social research methods, primarily configurational ones such as Coincidence Analysis, Event Structure Analysis and Qualitative Comparative Analysis. He has taught nationally and internationally on QCA, and has published methodological and applied work in numerous journals, including the American Journal of Evaluation, Comparative Political Studies, Political Analysis, Sociological Methodology and Sociological Methods & Research. He is the author of the QCApro package, and a co-author of both the QCA package and the cna package for the R environment. For further biographical details, see his Homepage or his ResearchGate website.

3 Detailed Workshop Schedule (preliminary)

Each day is divided into four modules, with each module lasting 90 minutes. A consultation session is offered after the fourth module on days 2 and 3. All required and supplementary readings will be made available to registered students three weeks in advance. Main readings in either category, required and supplementary, are marked with a “•” sign. A “+” sign indicates additional readings that closely relate to the respective main reading below which it is listed.
<table>
<thead>
<tr>
<th>Day</th>
<th>Module and Topic(s) Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Day 1: Monday, 25 September 2017: Theoretical Foundations</strong>&lt;br&gt;<strong>Detailed schedule</strong></td>
</tr>
<tr>
<td>09:00 - 10:30</td>
<td><strong>Module 1.1:</strong> Theorizing about causation and the essentials of Boolean algebra</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
<tr>
<td>10:45 - 12:15</td>
<td><strong>Module 1.2:</strong> Regularity theories from Hume over Mill to Mackie</td>
</tr>
<tr>
<td>12:15 - 13:30</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>13:30 - 15:00</td>
<td><strong>Module 1.3:</strong> Discovering regularity theoretic causation</td>
</tr>
<tr>
<td>15:00 - 15:15</td>
<td>Break</td>
</tr>
<tr>
<td>15:15 - 16:45</td>
<td><strong>Module 1.4:</strong> The basic work flow of QCA</td>
</tr>
</tbody>
</table>

**Required readings**

**Supplementary readings**

<table>
<thead>
<tr>
<th>Day</th>
<th>Module and Topic(s) Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Day 2: Tuesday, 26 September 2017: Crisp-Set QCA</strong>&lt;br&gt;<strong>Detailed schedule</strong></td>
</tr>
<tr>
<td>09:00 - 10:30</td>
<td><strong>Module 2.1:</strong> A short introduction to R</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
<tr>
<td>10:45 - 12:15</td>
<td><strong>Module 2.2:</strong> A short introduction to R</td>
</tr>
<tr>
<td>12:15 - 13:30</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>13:30 - 15:00</td>
<td><strong>Module 2.3:</strong> From raw data to the QCA solution</td>
</tr>
<tr>
<td>15:00 - 15:15</td>
<td>Break</td>
</tr>
<tr>
<td>15:15 - 16:45</td>
<td><strong>Module 2.4:</strong> Measures of fit in QCA</td>
</tr>
<tr>
<td>17:00 - 18:00</td>
<td>Consultation session</td>
</tr>
</tbody>
</table>
**Required readings**


**Supplementary readings**


---

**Day 3: Wednesday, 27 September 2017: Other variants of QCA**

<table>
<thead>
<tr>
<th>Time</th>
<th>Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 - 10:30</td>
<td><strong>Module 3.1:</strong> The Three Solution Types of QCA</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
<tr>
<td>10:45 - 12:15</td>
<td><strong>Module 3.2:</strong> Multi-Value QCA</td>
</tr>
<tr>
<td>12:15 - 13:30</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>13:30 - 15:00</td>
<td><strong>Module 3.3:</strong> Fuzzy-Set Theory and Fuzzy Logic</td>
</tr>
<tr>
<td>15:00 - 15:15</td>
<td>Break</td>
</tr>
<tr>
<td>15:15 - 16:45</td>
<td><strong>Module 3.4:</strong> Fuzzy-Set QCA</td>
</tr>
<tr>
<td>17:00 - 18:00</td>
<td>Consultation session</td>
</tr>
</tbody>
</table>
**Required readings**


**Supplementary readings**


**Day 4: Thursday, 28 September 2017: Critiques of QCA**

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 - 10:30</td>
<td>Module 4.1: A closer look at model ambiguities</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
<tr>
<td>10:45 - 12:15</td>
<td>Module 4.2: Criticisms of QCA</td>
</tr>
<tr>
<td>12:15 - 13:30</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>13:30 - 15:00</td>
<td>Module 4.3: QCA and mixed/multi-method research</td>
</tr>
<tr>
<td>15:00 - 15:15</td>
<td>Break</td>
</tr>
<tr>
<td>15:15 - 16:45</td>
<td>Module 4.4: Beyond QCA: causal chains and common cause structures</td>
</tr>
<tr>
<td>17:00 - 18:00</td>
<td>Consultation session</td>
</tr>
</tbody>
</table>
Required readings


Supplementary readings


Day 5; Friday, 29 September 2017: Coincidence Analysis  
* * *  detailed schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 - 10:30</td>
<td>Module 5.1: The CNA algorithm</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
</tbody>
</table>
Module 5.2: The differences and commonalities of QCA and CNA

Module 5.3: Working with the cna package for R

Module 5.4: The causal chain problem

Required readings

Supplementary readings
- Baumgartner, Michael, and Alrik Thiem. 2015. Identifying complex causal dependencies in configurational data with Coincidence Analysis.” The R Journal 7 (1).

4 Detailed Seminar Schedule

Day 1: Monday, 25 September 2017

The first module of the seminar introduces the basic objectives and designs of theories of causation, discusses various theory candidates, and presents the elements of Boolean algebra implemented by QCA and CNA. Module 1.2 then displays the details of the regularity theory of causation behind configurational methods and tracks its historical predecessors in the works of David Hume, John Stuart Mill, and John L. Mackie. In module 1.3 we turn to the problem of discovering causation as defined by modern regularity theories. We review Mill’s famous method of difference, consider the problems of causal inference under epistemic uncertainty and of data confounding, and pinpoint the background assumptions required by configurational methods. Finally, module 1.4 introduces the basic procedural ideas that regulate QCA’s inference from configurational data via truth tables to minimized Boolean functions (solutions formulas) and their causal interpretation. In particular, we present the details of QCA’s algorithmic core: Quine-McCluskey optimization.
**Day 2: Tuesday, 26 September 2017**

Day 2 begins with a short introduction to the R environment and its basic functionality in modules 2.1 and 2.2. In module 2.3, participants recap the material from module 1.4 by performing their first crisp-set QCA in an ideal data context from beginning to end with the QCApro package for R. Important differences between the QCApro package and popular QCA software like fs/QCA and Tosmana are also highlighted. Module 2.4 then introduces measures of fit in QCA for evaluating certain data deficiencies that are ubiquitous in applied research configurational data. In this connection, students learn about the consistency and coverage measures, and limited empirical diversity—the fact that certain cases which could theoretically exist are absent in the data.

**Day 3: Wednesday, 27 September 2017**

Much of the methodological literature of the past years has focused on the problem of limited empirical diversity, and how to deal with it. As a direct result, three solution types are now in use with QCA. Module 3.1 presents the logic of these three different solution types, and explains why only one of them, namely the parsimonious solution, is correct. Since about the early 2000s, QCA has also become a family of different variants that are defined by their underlying set type. Crisp-set QCA has at its root bivalent factors whose levels underlie sets in which cases can only be members or not. Both multi-value and fuzzy-set QCA extend crisp-set QCA in different directions, the former on the dimension of the number of levels a factor can have, and the latter on the dimension of the degree to which cases can be elements in the respective set formed by each level of a bivalent factor. In module 3.2, students learn about multi-value QCA, how it relates to crisp-set QCA, and why it still leads a niche existence in both methodological and applied research. Fuzzy logic and fuzzy-set QCA is the topic of modules 3.3 and 3.4. Just as on day 2, day 3 is structured around alternating theoretical and practical slots in which students directly apply the theoretical material in computer exercises.

**Day 4: Thursday, 28 September 2017**

Critiques of QCA are the core topic of day 4. In module 4.1, a closer look is first taken at the issue of model ambiguities, a problem that has gone unnoticed in the QCA literature until very recently, with serious consequences for applied research. We show why and how this problem occurs, to which extent it affects applied research, and what can be done to alleviate it. In module 4.2, we confront prominent studies which have argued, for various reasons, that QCA is useless. We demonstrate why these studies lack methodological traction. To what extent QCA is integrated and combined in mixed/multi-method research with regres-
sion analysis and process tracing is discussed in module 4.3. The last module on day 4 prepares students for day 5. In module 4.4, we demonstrate that the restriction of QCA to single outcomes presupposes that there are no causal dependencies among the exogenous factors.

**Day 5: Friday, 29 September 2017**

The aim of the final day is to pave a way for overcoming QCA’s restriction to single-outcome structures. After all, causes of ultimate outcomes being linked in chains or causes having multiple parallel outcomes are very frequent in the world we live in. To uncover such structures, day 5 introduces Coincidence Analysis (CNA). Module 5.1 presents the algorithmic protocol of CNA, reviews its theoretical foundation, and introduces complex solutions formulas. In module 5.2, we highlight the differences and commonalities of QCA and CNA: both methods analyze the same type of data and have the same search targets, but while QCA implements Quine-McCluskey optimization, CNA draws on its own custom-built optimization routine, which does not force CNA to make recourse to counterfactual reasoning in cases of limited diversity and does not require an outcome specification as input. Module 5.3 is then devoted to acquiring familiarity with the cna package for R and to giving students the opportunity to explore possibilities of making use of CNA in their own research. The seminar ends with module 5.4 presenting the causal chain problem, in virtue of which to every causal chain there exists a common cause model that is empirically indistinguishable from the chain. That means the inference to causal chains is systematically underdetermined by empirical data.

5  **Prerequisite Knowledge**

The course requires no prior formal knowledge, but it will be intensive for absolute beginners. Users with an intermediate to advanced knowledge of QCA as taught in standard textbooks and methods courses will learn much they did not know before. Participants at all levels of knowledge will benefit from the instructors’ current research, a significant part of which will be broached during the course. Basic knowledge of R, or at least programming more generally, will be helpful but is not essential.

6  **List of QCA Studies in Health Sciences**


41. Spangaro, Jo, Sigrid Herring, Jane Koziol-McLain, Alison Rutherford, Mary-Anne Frail, and Anthony B. Zwi. 2016. “They aren’t really black fellas but they are easy to talk to: Factors which influence Australian Aboriginal women’s decision to disclose intimate partner violence during pregnancy.” *Midwifery* 41:79-88.


