Using Mixed-Methods Evaluation Methods Taking Into Account Gender/Class Realities: Using QCA and NVIVO
Wendy Olsen
Funded by British Academy: Innovation in Global Labour Research Using Deep Linkage and Mixed Methods
Applications to:

- Theorising Bangladesh Indebtedness
  - Mediated by involvement in an NGO
  - All NGOs are different; and

- Indian Women’s Work
  - Mediated by involvement in either the self-help groups, an MFI,
    Or an NGO, or the Employment Guarantee Scheme
Steps for a Mixed-Methods Evaluation Approach

- Step 1: a complex theory of the ontic realities, i.e., the types of things
- Step 2: fieldwork
- Step 3: analysing early, & linking results
- Step 4: keyness, discourses
- Step 5: perhaps QCA analysis
- Step 6: transparency: database
- Step 7: draw conclusions
Key Sampling Themes

- **Representativeness** at some level
- Idea of the *replication of entities* across a geographic space
- **Generalisation** to known sub-populations and concrete spaces
Step 1: a complex theory of the ontic realities, ie the types of things

☐ The ontic reality is treated by statisticians as Structured
  ☐ Outcome = result of structures, events.

Logic is
  ☐ Y = results arise from S, I, E, C, random error

I = institutions, local entities

C = context

A non-statistical approach.
Discussion of Key Sampling Themes
ADVICE

- You may **triangulate** a national dataset onto your local data
  - **Match questions on demographics**, take a random sample not non-random!
  - Randomness at some, not all levels is, overall, non-random
  - But generalisation can be made at the level-to-which randomness was applied, e.g. by geographic transect walks.
    - E.g. a village. Or all the Slums of Dhaka if the Slums were stage 1 and
    - The choice of households was stage 2
    - And the choice of individuals (KISH) was stage 3
    - So be very professional about selection of cases.
  - **Multi-stage quota sampling vs. Multi-stage RANDOM sampling**: The difference is in the degree of REPLACEMENT of non-response cases.
You may **triangulate** a national dataset onto your local data

- **Match questions on demographics**, take a random sample not non-random!
- Randomness at some, not all levels is, overall, non-random
- But generalisation can be made at the level-to-which randomness was applied, e.g. by geographic transect walks.
  - E.g. a village. Or all the slums of Dhaka. The slums were stage 1 and
  - The choice of households was stage 2
  - And the choice of individuals (KISH) was stage 3
- So be very professional about selection of cases.

Multi-stage quota sampling vs. Multi-stage RANDOM sampling: The difference is in the degree of replacement of non-response cases.

Never sample on the outcome variable. For example on income levels, if you want to explain the change in income over time.
NVIVO Keyness Analysis of Discourses in Large Dataset (With Example of Matrix Results from South India)

Wendy Olsen 2016

Gender Norms Project

Acknowledging funding of ESRC DFID Pov. Allev. Fund and British Academy
Next Steps:

- Step 2: fieldwork
- Step 3: analysing early, & linking results
- Step 4: keyness
# How to conduct a Keyness Analysis for a Social Science Research Project.

1. Pool all the transcripts
2. Find out the keyness of words
3. Code up the concordances
4. Group the words into discourses
5. Interpret selected discourses only
6. Treat each one of those very carefully: the dominant discourse must be discerned, then the marginalised, deviant and innovative (intertextual) ones.
7. Trace key arguments through these. (Mixed Methods)
How the Keyness Analysis is Done

1. Keyness of words
2. Discourses too
3. Interpretations: dominant discourse; Marginalised & intertextual ones.
4. Trace key arguments through these. (Mixed Methods)

Key References:
- Fairclough, Norman various, books on *Discourse and Power*. 
Part One: The Keyness of Words
(Touri and Koteyko 2014)

Keyness is the relative prevalence of words in one corpus of material over another.

Specifically, count $S$ words in corpus, vs. $N$ words in the British National Corpus of English Language.

Use the formula provided here.
**Formula for Keyness**

- *Keyness* = odds ratio
- The odds of a word appearing in the fieldwork based corpus vs. the odds of it appearing in the national corpus

\[ K = \frac{s_i}{S-s_i} \frac{n_i}{N-n_i} \]  

For each word I

Counting words using NVIVO then

Matching words using STATA or SPSS

Report output as a word list, RANKED.
Word Count Query in NVIVO

- Look for: [Input]
- Search In: [Options]
- Queries: [Search terms]

- Word Frequency Criteria:
  - Search in: [All Sources, Selected Folders]
  - Display words: [Most frequent, All]
  - With minimum length: [Number]

- Results:
  - Word | Length | Count | Weighted Percentage (%)
  - group | 5     | 209   | 0.46
  - loans | 5     | 195   | 0.43
  - credit | 6     | 185   | 0.41
  - groups | 6     | 179   | 0.39
  - services | 8     | 177   | 0.39
  - financial | 9     | 169   | 0.37
  - company | 7     | 150   | 0.35
  - members | 7     | 140   | 0.33
  - women | 5     | 151   | 0.33
  - development | 11 | 146   | 0.32
  - basis | 5     | 130   | 0.30
  - crore | 5     | 121   | 0.27
  - [...]

- Run Query | Add to Project
Delemmatised of mentions (mention) Word Length Count Percent BNC Prevalence BNC % of BNC Odds
brickfields 11 2 0% 2 0% 4738.06
laws' 5 2 0% 3 0% 3158.71
purdah' 7 1 0% 2 0% 2369.03
coops 5 2 0% 11 0% 861.47
passbook 8 1 0% 6 0% 789.68
betel 5 3 0% 23 0% 618.01
mindset 7 1 0% 11 0% 430.73
parishad 8 2 0% 25 0% 379.04
stipends 8 2 0% 38 0% 249.37
negatively 10 2 0% 52 0% 182.23
sons' 5 1 0% 28 0% 169.22
educate 7 12 1% 365 0% 155.77
workloads 9 1 0% 43 0% 110.19
rears 5 1 0% 45 0% 105.29
chores 6 6 1% 275 0% 103.38
robbers 7 5 0% 244 0% 97.09
tailoring 9 3 0% 147 0% 96.70
dhaka 5 1 0% 55 0% 86.15
We created a spreadsheet to rank words by their keyness ratio. Here is the table:

### Table A1: Entire list of 233 high keyness words

<table>
<thead>
<tr>
<th>Word</th>
<th>Count in SSI</th>
<th>Count in Baby BNC</th>
</tr>
</thead>
<tbody>
<tr>
<td>10000</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>15000</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>20000</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>200ft</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>30000</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>350ft</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>40000</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>50000</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>500ft</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>acres</td>
<td>174</td>
<td>35</td>
</tr>
<tr>
<td>adjoining</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>agreement</td>
<td>67</td>
<td>232</td>
</tr>
<tr>
<td>agriculture</td>
<td>151</td>
<td>121</td>
</tr>
<tr>
<td>alias ? !</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>allah</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>anymore</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>approx</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>artisan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crusher</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crushing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cultivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>daughter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>departure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>depth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>devotee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dignity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>disagreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>disrespect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>doubts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dowry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>draught</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drought</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eldest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eligible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>entrusted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>erstwhile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>expenditure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explicit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>extras</td>
<td></td>
<td></td>
</tr>
<tr>
<td>family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In this example from South India, 39 interviews with couples.

- 39 Interviews
- 47,000 Words
- We reduced these to 233 key words. **Extremely concise summary.**
- Then as an expert I examined these to group them into discourse topics.
- Next I study these discourse topics to identify discursive patterns.
Interim Product Conforms to Miles & Huberman’s Advised “one-page summary”
Here’s an example (a small South Indian project)

APPENDIX 1: extra tables to illustrate matters from South Indian Mixed Methods Tenancy Project

Table a1: Entire list of 233 high keyness matched words from the SSIs (Alphabetical)

<table>
<thead>
<tr>
<th>Word</th>
<th>Count in SSI</th>
<th>Count in Baby BNC</th>
<th>odds ratio BNC</th>
<th>Size of Baby BNC</th>
<th>Size of SSI Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>10000</td>
<td>10</td>
<td>6</td>
<td>0.0002</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>15000</td>
<td>2</td>
<td>1</td>
<td>0.00009</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>20000</td>
<td>2</td>
<td>1</td>
<td>0.0001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>200ft</td>
<td>2</td>
<td>5</td>
<td>0.00001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>30000</td>
<td>5</td>
<td>2</td>
<td>0.00001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>350ft</td>
<td>1</td>
<td>1</td>
<td>0.0000</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>40000</td>
<td>2</td>
<td>1</td>
<td>0.0000</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>50000</td>
<td>5</td>
<td>2</td>
<td>0.00001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>500ft</td>
<td>1</td>
<td>2</td>
<td>0.0001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>acres</td>
<td>174</td>
<td>35</td>
<td>0.0009</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>adjoining</td>
<td>3</td>
<td>19</td>
<td>0.0005</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>agreement</td>
<td>67</td>
<td>232</td>
<td>0.0059</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>agriculture</td>
<td>151</td>
<td>121</td>
<td>0.0031</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>alias</td>
<td>153</td>
<td>3</td>
<td>0.0001</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>allah</td>
<td>3</td>
<td>1</td>
<td>0.0000</td>
<td>39701</td>
<td>2741</td>
</tr>
<tr>
<td>approx</td>
<td>20</td>
<td>10</td>
<td>0.0003</td>
<td>39701</td>
<td>2741</td>
</tr>
</tbody>
</table>
Annotate and summarise the Key Terms.
Group them into dominant discourses.
This is also like thematic analysis, initially.

Discourses are sets of rules which are coherent but which are held to only via normed practices, and which can be broken, at a certain price.

Example of patriarchal talk about marriage as an exchange of assets.

Next: Locate the marginalised discourses
Steps for a Mixed-Methods Evaluation Approach

- **Step 1**: a complex theory of the ontic realities, i.e., the types of things
- **Step 2**: fieldwork
- **Step 3**: analysing early, & linking results
- **Step 4**: keyness, *discourses*
- **Step 5**: perhaps QCA analysis
- **Step 6**: transparency: database
- **Step 7**: draw conclusions
Discourses we found (South India; North India)

- **Dominant ones:**
  - Agriculture as production
  - Family as duty, obligations (disciplining)
  - Moneylending as a solution

- **Marginal ones:**
  - Agriculture as a burden the older generation carry, disliked
  - Family as conflict
  - Moneylending and debt as a problem
**SCALE of the DATABASE: A Small Research Project in Bangladesh**

<table>
<thead>
<tr>
<th>Interviews</th>
<th>Words</th>
<th>Stemmed Words</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 interview</td>
<td>673 raw words of 5+ letters</td>
<td>396 “words” i.e. word-roots, in one interview, if you stem the words</td>
<td>By hand</td>
</tr>
<tr>
<td>11 interviews</td>
<td>1666 words</td>
<td>1249 after stemming</td>
<td>By NVIVO</td>
</tr>
<tr>
<td>32 interviews</td>
<td>2798 words</td>
<td>2066 word-roots, after stemming</td>
<td>By NVIVO</td>
</tr>
</tbody>
</table>
Using the Words with Highest Keyness

- We set a cutoff level for keyness (the odds ratio) e.g. 4, or 9.
- Collect the concordances using NVIVO
- You now have extensive quotations to compare and contrast.
- Link the survey data to this database.
REMINDER: My Keyness Method

1. Pool all the transcripts using NVIVO.

2. Code up the concordances

3. Group the words into discourses

4. Interpret selected discourses only

5. Treat each one of those very carefully

6. Trace key arguments through them.
COMPARATIVE NVIVO

Results for two discourses (family talk and money talk) [india 1 and bangla 1 combined] Mentioned within 30 words of each other, in combination.

<table>
<thead>
<tr>
<th></th>
<th>A: Tightness node</th>
<th>B: money</th>
<th>C: problems</th>
<th>D: spend</th>
<th>E: works</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Family</td>
<td>20</td>
<td>18</td>
<td>11</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>2: children</td>
<td>12</td>
<td>10</td>
<td>6</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>3: daughter</td>
<td>15</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>4: husband</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>5: mother</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>11</td>
</tr>
</tbody>
</table>
Steps for a Mixed-Methods Evaluation Approach

- Step 1: a complex theory of the ontic realities, ie the types of things
- Step 2: fieldwork
- Step 3: analysing early, & linking results
- Step 4: keyness, discourses
- **Step 5: perhaps QCA analysis**
- Step 6: transparency: database
- Step 7: draw conclusions
Qualitative Comparative Analysis

Logic is

- \( Y = \) results arise from \( S, I, E, C, \) random error

\( I = \) institutions, local entities

\( C = \) context

A non-statistical approach.

Is event \( E \) necessary, or sufficient for \( Y \)?
Aims and Means of QCA

Aims
- To focus on one outcome.
- How does the effect of X or T or E on that outcome change depending upon the contexts?
- Circumstances matter.
- Measure to what extent it was the case.

Means
- Insert a survey matrix into fsQCA freeware.
- Produce tests of necessity of EACH condition for Y.
- Then test for sufficient PATHWAYS.
- Test the results using a measure, or an F Test.
- See my GITHUB freeware.
Details of the QCA F-Tests

1 We first define our terms and conceptual framework (S, I, E, X, Y, C)

2 **Empirical measure of Csuff** (consistency for sufficiency of X for Y)

3 **Empirical measure of Goodness-of-fit (F-tests)** for each pathway to Y

See
https://github.com/WendyOlsen/fsgof
Amending the QCA for treatments, impacts of interventions

- In logic add ‘T’ as a new event
- Allow it to work as a ‘necessary’ cause (test) of higher levels of Y
- Allow it to be considered as a sufficient pathway for higher levels of Y
- Allow it to be considered as part of sufficient combination pathways for higher levels of Y
Practical Example
Applications to:

- Theorising Bangladesh Indebtedness
  - Mediated by involvement in an NGO
  - All NGOs are different; and

- Indian Women’s Work
  - Mediated by involvement in either the self-help groups, an MFI, Or an NGO, or the Employment Guarantee Scheme
Sample of Raw Debt Data – Bangladesh
Results from QCA Part for India
Reminder: **Mixed Mode Data**

- **Step 1:** ontic exploration, list the types of things, name the key processes,
  - SAMPLING: Get samples which have CONTRASTS on BOTH X and Y
  - AND ON T, the treatment event (low/high!) or (Yes/No)
  - And on contextual factors (see leaflet)
  - Make sure the qualitative cases are chosen from among the pre- and post-intervention sample cases.
- **Step 2:** fieldwork
- **Step 3:** analysing early, & linking results
- **Step 4:** keyness, discourses
- **Step 5:** perhaps QCA analysis
- **Step 6:** transparency: database
- **Step 7:** draw conclusions
Discussion
Critiques and Responses

- RCT critique
- Unobserved heterogeneity critique

Responses: Complex differentiation of how causal mechanisms work
Endogeneity critique

(it says that the key factors in your model can’t be distinguished from the irrelevant ones you have included because you’ve included too many factors)

Responses:

Complex interactions $\Rightarrow$ do not ignore possible pathway reversal phenomena!

That’s why statistics is weaker.

Furthermore, be parsimonious in setting up the QCA explanatory model.
Conclusions

- Ontic complexity
- Teamwork
- Combining the keyness stage with a selective interpretation stage; and
- Add A QCA or Fuzzy Set QCA Stage.
- Models and results are debated in an ongoing, open-ended way.
- We try to make the interpretation match, complement or contradict the original Research Question.
- Be rigorous and transparent.
Acknowledgements - Collaborators **John McLoughlin** at Univ of Manchester, **Samantha Watson** at Flowminder, University of Southampton

- John has programmed in Python to break up the British National Corpus into parts and put them into NVIVO. Counting the word frequencies in Baby BNC in NVIVO, he then compared these with the word frequencies in each qualitative data set.
- See GITHUB for the programme, searching on either John McLoughlin or Wendy Olsen.
See Also:

- See also a calibration example at: https://www.facebook.com/groups/mixednetwork/
- Integrated Mixed Methods Network
- And many examples of QCA and Fuzzy Set Analysis of Cases at www.compasss.org (sic)
- And JISCMAIL QUAL-COMpare (190 members) email list. Free to join.
Key References


