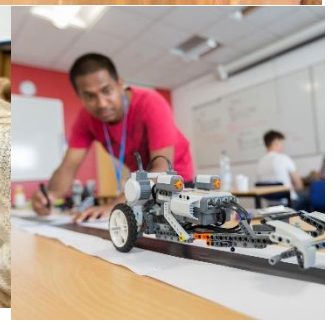
NERUPI *network*
Evaluating & Researching University Participation Interventions

Comparative Evaluations

19th October 2023

Joanne Moore, NERUPI Development Officer, University of Bath

Annette Hayton, NERUPI Convenor. Senior Research Fellow, University of Bath



Approaches to comparisons and causality

NERUPI network
Evaluating & Researching University Participation Interventions

1. **Quantitative and Quasi-Scientific approaches**
   - Changes compared to baseline (participants)
   - Comparison of changes for two groups of participants in controlled environments (e.g. outcomes different locations/delivery models)
   - Changes for participants against a control group (random experiment) or comparison group (quasi-experimental methods)
2. **Qualitative and Realist approaches**
   - Focus on the context and theory of practices
   - What works for whom in what circumstances

The Office for Students' Standards of Evidence (OfS, 2019)
Three types: Narrative, Empirical, Causal

Plan for the session

NERUPI network
Evaluating & Researching University Participation Interventions

1. **Basis for comparisons and making causal inferences**
2. **Quantitative approaches**
   - When to use
   - Examples of different types of comparisons and groups
   - Different analytical techniques
3. **Qualitative approaches**
   - When to use
   - Example of Qualitative Comparative Analysis (QCA)
   - Sign-posting
4. **Concluding remarks and questions**

John Stuart Mill's Five 'Canons'

NERUPI network
Evaluating & Researching University Participation Interventions

The 'Canon'	Experiment	Finding
1. Method of agreement	Looks at a factor in common, with the same outcome	"*If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree is the cause (or effect) of the given phenomenon"*
2. Method of difference	Looks at the outcome in the absence of one factor	*"If every circumstance [is] in common save one… the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon"*
3. Indirect method of difference	**Applies the method of agreement *before and after* an event or external stimulus**	**The event/stimulus is seen as responsible for the changed outcome**
4. Method of residues	Repeated use of the method of agreement	Elimination of previous causes
5. Method of concomitant variation	Correlation between two factors and a common external factor	The outcome is dependent on the external factor which can't be eliminated

Ref: John Stuart Mill Chapter VIII

Activity 1: Pop's Seafood Diner

NERUPI network
Evaluating & Researching University Participation Interventions

Four people ate out on their way home, they all became sick….

1. First Person - Tasted a little bit of everything
2. Second Person – Oysters and Halibut
3. Third Person – Only Oysters
4. Fourth Person – Oysters and Shrimps

- **Which food caused the illness?**
- **Vote now!**

https://www.menti.com/alic54vkwpx3

Moses and Knutsen (2007), Ways of Knowing: Competing Methodologies in Social and Political Research, London: Bloomsbury Press

Join at menti.com use code 4449 096

Which food caused the illnes

Activity 1: Pop's Seafood Diner

NERUPI network
Evaluating & Researching University Participation Interventions

Four people ate out on their way home, they all became sick

1. First Person - Tasted a little bit of everything
2. Second Person – Oysters and Catfish
3. Third Person – Only Oysters
4. Fourth Person – Oysters and Shrimps

Plus 3 more people who didn't become sick

5. **Fifth Person – Only Shrimps**
6. **Sixth Person – Only Halibut**
7. **Seventh Person – Shrimps and Haibut**

Indirect method of difference – more certainty that oysters were responsible

Considerations

NERUPI network
Evaluating & Researching University Participation Interventions

<table>
<tr><td></td><td colspan="4">Number of Cases</td></tr>
<tr><td rowspan="4">Number of Variables</td><td></td><td>1-2</td><td>Small N</td><td>Medium N</td><td>Large N</td></tr>
<tr><td>Large V</td><td rowspan="2">Paired comparison</td><td>Comparative methods like QCA</td><td rowspan="2">Quantitative Methods</td><td rowspan="2">Multi-variate Statistical Analysis</td></tr>
<tr><td>Small V</td><td></td></tr>
<tr><td>1-2</td><td>Description</td><td colspan="3">Bivariate descriptive classification</td></tr>
</table>

V=variables; N=cases

Multiple dimensions

NERUPI network
Evaluating & Researching University Participation Interventions

- **Objective dimension**
  - Organisations, social structures, institutions
- **Subjective dimension**
  - Perceptions of the world, sentiments, beliefs
- **Normative dimension**
  - Principles of human lives, values, judgements

Behavioural outcomes

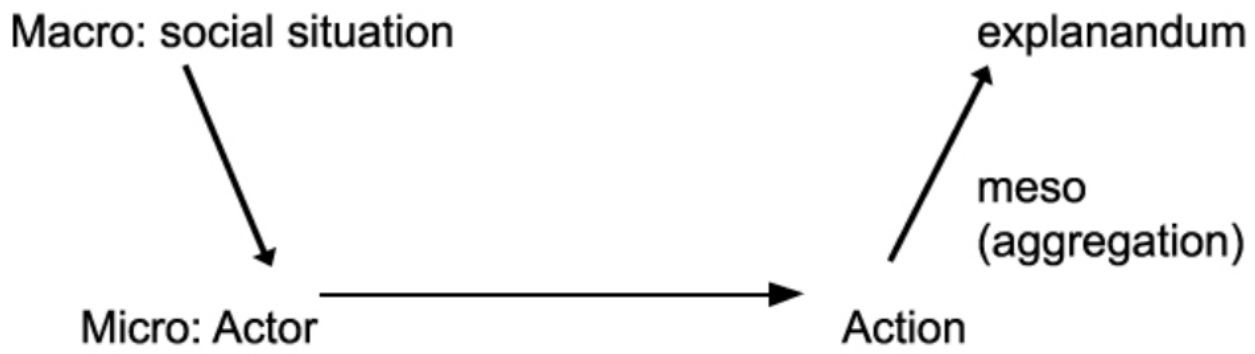
Actions

Non-behavioural outcomes

Thoughts, perceptions, feelings

Multiple dimensions

NERUPI network
Evaluating & Researching University Participation Interventions



Adapted from Coleman 1990 and Esser 1993

Macro level - societies, communities

Meso level – organisations, interest groups

Micro level - individual persons and actors

Quantitative approaches

NERUPI network
Evaluating & Researching University Participation Interventions

Orientation	Variable orientated
Explanatory Focus	• Effects of causes
Concepts	• Linear, cumulative • Independent effect of each variable
Measurement	• Quantitative - Cross case comparison - Estimation of average effects
Cases	• Large N - All cases/Random/Systematic sampling
Aims	• Proof of causation/promising results • Estimation/Probability/Generalisations

Types of data/outcome measures

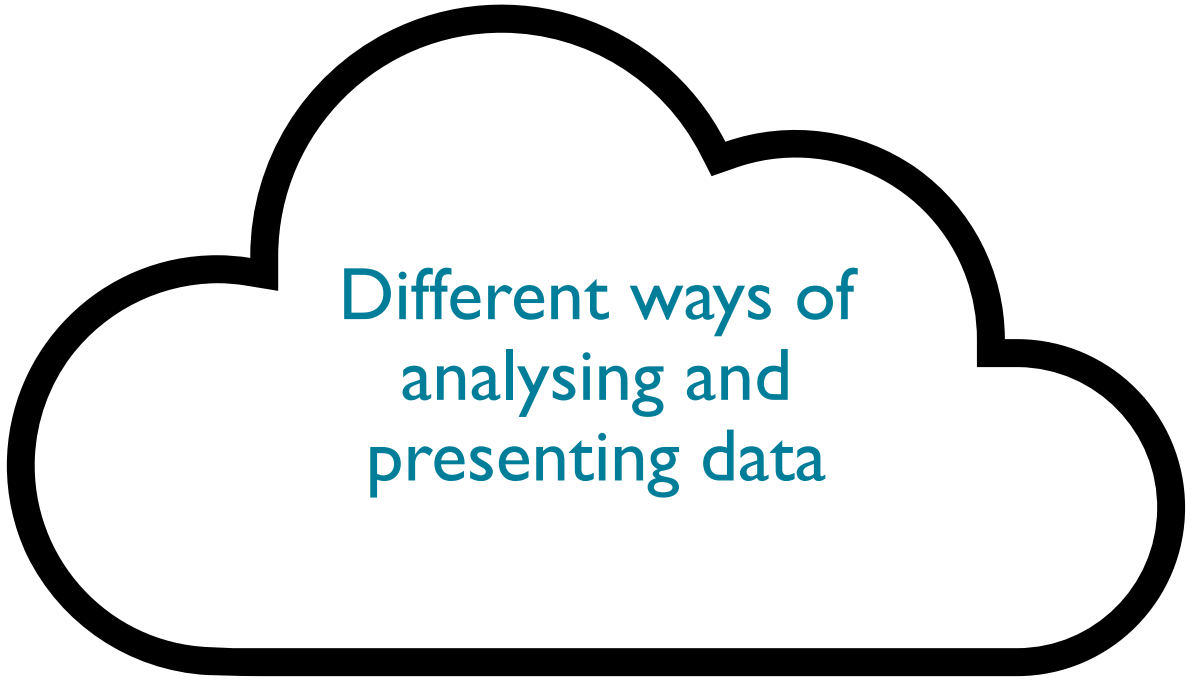
NERUPI network
Evaluating & Researching University Participation Interventions

- Assessments/tests/performance (before, after intervention)
- Decision making (e.g. applications to HE)
- Tracking data (e.g. applications by type of HEI, by course)
- Attainment/exams (e.g. school, university course results)
- Survey question results (e.g. intentions, expectations, perceptions)

Analytical approaches

NERUPI network
Evaluating & Researching University Participation Interventions

- **Descriptive statistics**
  - Comparing means or looking at standard deviations
  - Tests of significance of differences and the effect size
  - Correlations and similar procedures
- **Statistical methods**
  - Regression analysis
  - Multivariate statistical analysis



Example 1: Before and After (individuals)

NERUPI network
Evaluating & Researching University Participation Interventions

https://www.nerupi.co.uk/public/assets/images/Evaluation-Wheel.pdf
https://www.nerupi.co.uk/members/resources/indicator-wheel

Example 2: Before and After (whole group)

NERUPI network
Evaluating & Researching University Participation Interventions

Table 1: Agreement with statements

A	B	C	D	E
Statement	% Pre* (N=103)	% Post* (N=95)	Percentage point Change	% of respondents with a positive change
People like me don't go to university (reverse scored)	30.1%	22.1%	8	27%
I feel well-prepared to make decisions about my next steps in education	59.6%	73.7%	14.1	24%
I have a good idea of what it's like to be a university student	92.2%	92.6%	0.4	0%
I understand how I can use what I'm learning in the future	76.7%	78.9%	2.1	3%
I think I have the skills I will need to be successful in my future studies	15.4%	20.0%	4.6	30%
Deciding what subject options are right for me feels overwhelming (reverse scored)	26.9%	9.5%	17.4	65%

https://www.nerupi.co.uk/public/assets/images/Options-for-Benchmarking.pdf

Statistical test for pre/post designs I

NERUPI network
Evaluating & Researching University Participation Interventions

- T Test of significance (e.g. P<0.05)

					T-Test Results for Equity of Means (Equal variances assumed)		
Statement		Mean	Std. Deviation		t	df	Sig. (2-tailed)
People like me don't go to university (reverse scored)	Pre Post	2.75 2.65	1.03 1.16		0.61	196	0.543
I feel well-prepared to make decisions about my next steps in education	Pre Post	3.19 2.73	1.15 1.23		2.759	197	0.006

https://www.nerupi.co.uk/public/assets/images/Options-for-Benchmarking.pdf

Statistical test for cross-group comparisons

NERUPI network
Evaluating & Researching University Participation Interventions

- Chi-Squared Test (e.g. Pearson Chi-Square test (χ^2) of 'goodness-of-fit')

Table 3: "People like me don't go to university" cross-tabulated with HE background

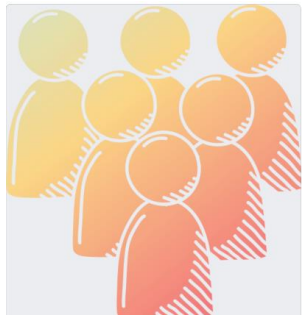
	Disagree	Agree	
No HE Background in Family			
Count	17	14	31
% within HE Background	54.8%	45.2%	100.0%
% within S1	25.0%	46.7%	31.6%
% of Total	17.3%	14.3%	31.6%
He Background in Family			
Count	51	16	67
% within HE Background	76.1%	23.9%	100.0%
% within S1	75.0%	53.3%	68.4%
% of Total	52.0%	16.3%	68.4%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4.509	1	.034

https://www.nerupi.co.uk/public/assets/images/Options-for-Benchmarking.pdf

NERUPI network
Evaluating & Researching University Participation Interventions



Discussion

Are these approaches useful?
Anyone using them?

Approaches to identifying control and comparator groups

	Method	Pros	Cons
Random selection	Randomly assign potential participants to either receive an intervention or 'business as usual' without the intervention	- Strong proof of causality - Different levels (individuals. Schools)	- Ethical, political and practical issues - 'Contamination' issues - Many results inconclusive
Propensity Score Matching (PSM)	Uses a statistically created comparison group based on an analysis of factors affecting propensity to participate	- statistically mitigates for selection bias	- Need the data to select comparators and access to their information. - Contamination issues
Regression Discontinuity Design (RDD)	Compares the outcomes of the intervention group with those just above/below the threshold	- Strength of the design depends on how similar the comparison group is to participants.	- Need data on those above/below threshold, e.g. through an application process
Difference in difference ('natural' experiment)	Before-and-after change in participants' relative to that of non-participants	- Better than simple before-after approach - Controls for 'common trends'	- Sample size and timeline challenges - Selection bias (not random)

https://www.nerupi.co.uk/members/resources/experimental-quasi-experimental-methods

HEAT KS4 comparator example (PSM)

NERUPI network
Evaluating & Researching University Participation Interventions

Average Attainment 8 Scores (Average grade across 8 subjects)

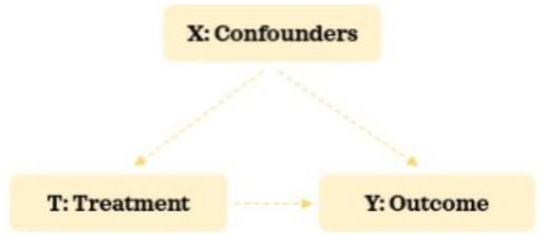
Average Attainment 8 Scores for low/average/high achievement band at KS2

	Participants	School Average	Difference
Low KS2 Attainment (<L4)	28.2	27.7	+0.5
Medium KS2 Attainment (L4)	44.7	40.3	+4.4
High KS2 Attainment (>L4)	59.0	54.0	+5.0

NB This means that on average medium attaining participants achieved an average of 4.4 grades higher when compared with pupils with similar attainment from th

Positively Mad Participants' Attainment 8 Scores compared with the Non-Participants' Scores

https://heat.ac.uk/impact-on-key-stage-4-attainment/



Difference-in-Difference (d-i-d) Example

NERUPI network
Evaluating & Researching University Participation Interventions

Ethnicity group	Outcome measures	Condition (reformed status)	Pre-intervention (average over 4 years) Mean (SD)	Post-intervention (average over 1~3 years) Mean (SD)	Descriptive difference-in-difference
BAME students	Module mark percentile rank	Comparator	38.9 (27.5)	39.4 (27.0)	(40.6-37.4) – (39.4-38.9) = **+ 2.7 percentiles**
		Reformed	37.4 (26.5)	40.6 (28.1)	
	% Achieving 2nd class and above	Comparator	38.8% (47.8%)	44.2% (49.7%)	(42.8-34.9) – (44.2-38.8) = **+ 2.5pp**
		Reformed	34.9% (47.7%)	42.8% (49.5%)	
White students	Module mark percentile rank	Comparator	53.5 (28.5)	53.7 (27.5)	(55.2-51.7) – (53.7-53.5) = **+ 3.3 percentiles**
		Reformed	51.7 (27.4)	55.2 (29.1)	
	% Achieving upper second class and above	Comparator	59.5% (49.1%)	65.3% (47.6%)	(62.7-57.3) – (65.3-59.5) = **– 0.4pp**
		Reformed	57.3% (49.5%)	62.7% (48.4%)	

TASO (2022) The impact of curriculum reform on the ethnicity degree awarding gap https://s33320.pcdn.co/wp-content/uploads/Full-report-the-impact-of-curriculum-reform-on-the-ethnicity-degree-awarding-gap.pdf

Regression Analysis

NERUPI network
Evaluating & Researching University Participation Interventions

Regression analysis

- relationship between a dependent variable and one or more independent variables.

Level 2/3 comparisons

- Need to control for the influence of confounding factors

Example

1. You have outcome data on summer school participants and on non-participants in a RCT on an over-subscribed summer school programme.
2. You calculate the average value of the HE applications of each of the two groups and compare them. The 'raw' difference in the mean outcome across the two groups shows the participants were more likely to say they'd make HE applications.
3. The result does not allow for the possible impact on outcomes associated with other things:
   - compositional differences across the two groups (e.g. sex, ethnicity)
   - characteristics (e.g. attainment) or engagement (e.g. in other outreach)
4. Regression analysis is needed to take account of the influence of observed confounding factors in estimating the impact of the Summer School.

RCT Example: Summer Schools

NERUPI network
Evaluating & Researching University Participation Interventions

Outcome	Treatment	Control
	Mean (SD)	Mean (SD)
Likelihood of going to HE (7-point Likert scale) (n = 342)	6.60 (0.99)	6.60 (0.98)
Likelihood of progressing to academic study post-16 (5-point Likert scale) (n = 49)	4.71 (0.52)	4.73 (0.46)
Self-efficacy relating to HE (5-point Likert scale) (n = 331)	4.06 (0.66)	3.91 (0.79)
Compatibility of HE with social identity (5-point Likert scale) (n = 337)	3.97 (0.95)	3.83 (0.97)
Perception of practical barriers to HE (5-point Likert scale) (n = 330)	3.38 (0.95)	3.26 (0.96)
Applied to HE (binary yes/no) (n = 295)	0.94 (0.23)	0.91 (0.29)
Notes: Sample of students (n) per outcome included in brackets above.		

https://taso.org.uk/news-item/new-report-interim-findings-on-the-impact-of-summer-schools-in-the-time-of-covid-19/

Effect Size

NERUPI network
Evaluating & Researching University Participation Interventions

Outcome	Estimated effect (score on scale)	Standard error	Estimated effect (Cohen's *d*)
Likelihood of going to HE (7-point Likert scale) (n = 342)	-0.01	0.11	-0.01
Likelihood of progressing to academic study post-16 (5-point Likert scale) (n = 49)	-0.12	0.17	-0.06
Self-efficacy relating to HE (5-point Likert scale) (n = 331)	0.14+	0.08	0.21
Compatibility of HE with social identity (5-point Likert scale) (n = 337)	0.15	0.11	0.14
Perception of practical barriers to HE (5-point Likert scale) (n = 330)	0.10	0.10	0.12
Applied to HE (binary yes/no) (n = 295)	0.04	0.03	0.14

Notes: N per outcome included in brackets above.

'Likelihood of going to HE' and 'Applied to HE' were computed for the post-16 sample only.

'Likelihood of progressing to academic study post-16' was computed for the pre-16 sample only.

All other effects were computed for the combined pre- and post-16 sample.

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

https://taso.org.uk/news-item/new-report-interim-findings-on-the-impact-of-summer-schools-in-the-time-of-covid-19/

Resources

NERUPI network
Evaluating & Researching University Participation Interventions

Guidance

Khandker, Koolwal & Sawad (2010) Handbook on Impact Evaluation: Quantitative Methods and Practices, Washington: World Bank

Statistical packages

- Excel Stats package
  Use the Analysis ToolPak to perform complex data analysis - Microsoft Support
- Introductory R and R Studio
  https://psyteachr.github.io/ads-v2/01-intro.html
- SPSS
  https://www.ibm.com/spss

Qualitative Comparative Analysis (QCA)

NERUPI network
Evaluating & Researching University Participation Interventions

Orientation	Case orientated
Explanatory Focus	• Effects of Causes
Concepts	• Explanation of specific cases • Necessary and sufficient causes • Multiple conjectural causation (different combinations)
Measurement	• Quantitative and qualitative - Limited variation
Cases	• Small-Medium N
Aims	• Understanding/reducing complexity • 'Middle range' contextualised generalisations

Implementation and process indicators?

NERUPI network
Evaluating & Researching University Participation Interventions

May be distinct and vary according to role/context

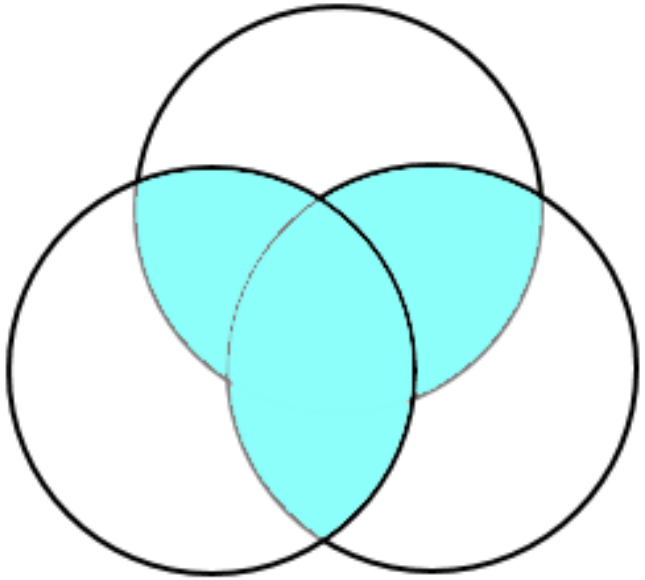
- Learning outcomes e.g. grades and marks
- Process outcomes e.g. attendance and engagement in the intervention or wider
- The learning environment e.g. delivery, curriculum
- The process of learning e.g. interaction with content

Analytical strategy

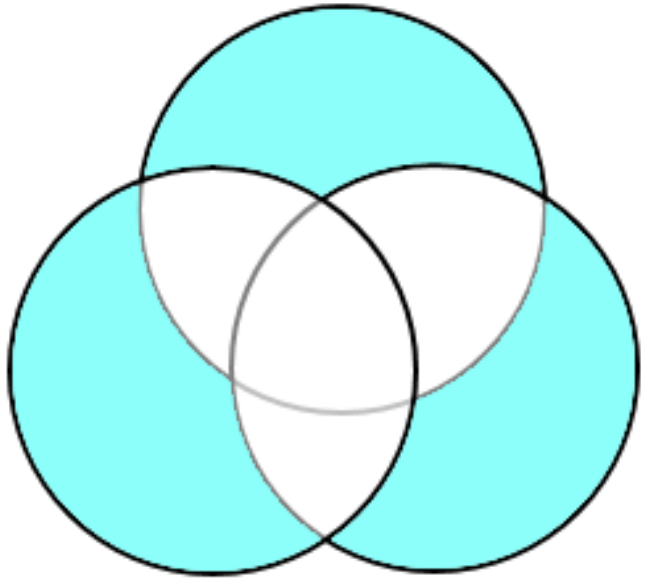
NERUPI network
Evaluating & Researching University Participation Interventions

- Qualitative comparative analysis (QCA)
  'Truth Table' (Boolean Algebra using specialist computer package)

Most Different Same Outcome (MDSO)



Most Similar Different Outcome (MSDO)



https://methods.sagepub.com/book/the-sage-handbook-of-case-based-methods/n12.xml

Steps

NERUPI network
Evaluating & Researching University Participation Interventions

1. Formulation of research question
2. Selection of (comparable) cases
3. Establishing similarities and differences:
4. Systematic matching of cases, MSDO
5. Systematic contrasting of cases, MDSO
6. Linking results to hypotheses
7. Further testing (e.g. with more cases, areas, time periods etc.)

https://methods.sagepub.com/book/the-sage-handbook-of-case-based-methods/n12.xml

QCA Truth Table Example

NERUPI network
Evaluating & Researching University Participation Interventions

Name	Girl (G)	Age 14yrs or older (A)	Same School (S)	Outcome (O)
Nicola	1	1	1	1
James	0	0	1	1
Jordan	0	1	0	0
John	0	1	1	1
Jayne	1	0	0	0
Sandra	1	0	1	1
Vicky	1	1	0	1
Paul	0	0	0	0

0 = Yes

1 = No

Rules:

- Pairwise
- Stepwise
- One reduction at a time

Results

NERUPI network
Evaluating & Researching University Participation Interventions

Variants

NERUPI network
Evaluating & Researching University Participation Interventions

- Binary data for both independent variables (conditions) and dependent variable (outcome) = crisp set (cs-) QCA
- Limited number of values (e.g. categorical data, nominal or ordinal scales) = multi-value (mv-) QCA
- Multi-value data for both conditions and outcomes = fuzzy set (fs-) QCA

	small N (~ 2-5 cases)	moderate N	larger N (>25 cases)
many variables	MSDO	MDSO cs-QCA mv-QCA	fs-QCA

Resources

NERUPI network
Evaluating & Researching University Participation Interventions

Guides

https://taso.org.uk/evidence/evaluation-guidance-resources/impact-evaluation-with-small-cohorts/what-is-small-n-evaluation/

https://www.nerupi.co.uk/public/assets/images/Qualitative-Comparative-Analysis.pdf

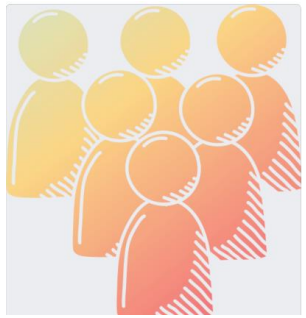
https://www.betterevaluation.org/methods-approaches/methods/qualitative-comparative-analysis

https://methods.sagepub.com/book/the-sage-handbook-of-case-based-methods/n12.xm

Software

- COMPASS
  https://compass.org/software
- QCA add in for excel
- TOSMANA
  https://www.tosmana.net/
- Also in the R software package

NERUPI network
Evaluating & Researching University Participation Interventions



Discussion

Are these approaches useful?
Anyone using them?

Some Concluding remarks

NERUPI network
Evaluating & Researching University Participation Interventions

Orientation	Variable orientated	Case orientated
Focus	• Effects of causes	• Effects of Causes
Concepts	• Linear, cumulative • Independent effect of each variable	• Explanation of specific cases • Necessary and sufficient causes • Multiple conjectural causation
Measurement	• Quantitative (cross case comparison, estimation of average effects)	• Quantitative and qualitative (categorisations)
Cases	• Medium - Large N	• Small-Medium N
Aims	• Proof of causation/results • Estimation/Probability/ Generalisations	• Understanding/reducing complexity • 'Middle range' contextualised generalisations

Areas for Development

NERUPI network
Evaluating & Researching University Participation Interventions

- Understanding linkages between short-medium-long term outcomes
  - How to define and measure the outcomes and impact?
  - Operationalisation of the intermediate learning outcomes which are useful predictors of longer term?
  - E.g. Application of validated scales
- Using Theories of Change (ToC) as the basis for setting up variables that can be used to capture effectiveness of different practice approaches

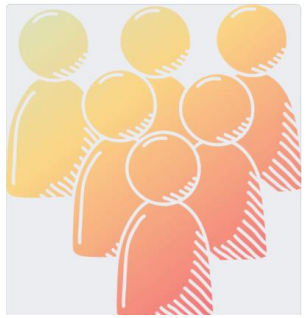
More info and suggestions for future topics?

NERUPI network
Evaluating & Researching University Participation Interventions

- What types of comparative evaluations are you doing at the moment?
- What more information would you like to assist with comparative evaluations?
- What other topics should we cover in the evaluation toolkit series?

https://padlet.com/nerupi/toolkit-requests-8mbks8sy9qsmluyt

NERUPI network
Evaluating & Researching University Participation Interventions



Thanks for Listening!

Questions and Answers

Multi-variate statistical analysis

NERUPI network
Evaluating & Researching University Participation Interventions

Regression analysis

- relationship between a dependent variable and one or more independent variables.

Analysis of Variance (ANOVA)

- relationship between collections of data by using the difference in the means

Cluster analysis

- similarities in a group of observations

Principal component analysis

- interpretation of uncorrelated variables

Factor analysis

- data reduction technique

Used to:

- understand patterns of data
- study multiple factors at once
- make clear comparisons
- discard unwanted information
- reduce complexity