THE UNIVERSITY OF WARWICK

2: Comparative analysis and case selection

Tina FREYBURG

[Introduction to Qualitative Methods]

Roadmap

Part I - **RESEARCH DESIGN**

- Causal thinking and research designs
- **Comparative analysis and case selection**

Part II - **DATA COLLECTION**

- (Semi-)Structured Interviews
- Observational research and ethical questions
- Observational research and ethical questions

Part III - **DATA ANALYSIS AND CAUSAL INFERENCE**

- Case studies and process-tracing
- Qualitative Comparative Analysis (QCA)
- The mixed-methods approach

Schedule

Part I: Research designs

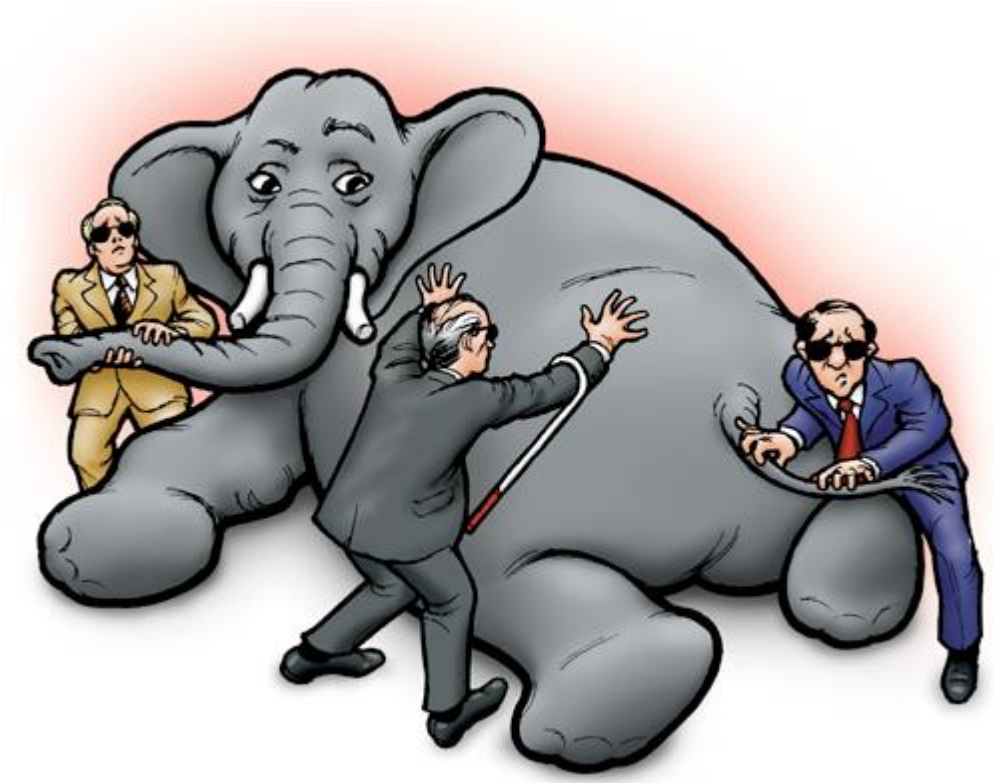
- Research design
- Your research project

Part II: Comparative analysis

- Comparisons
- Exercises

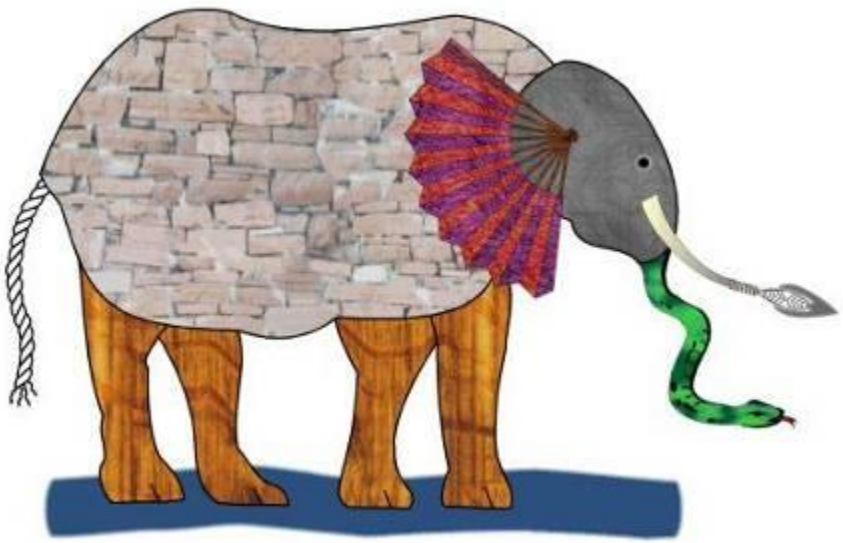
I -- The Parable of the Elephant

... and the Blind Men

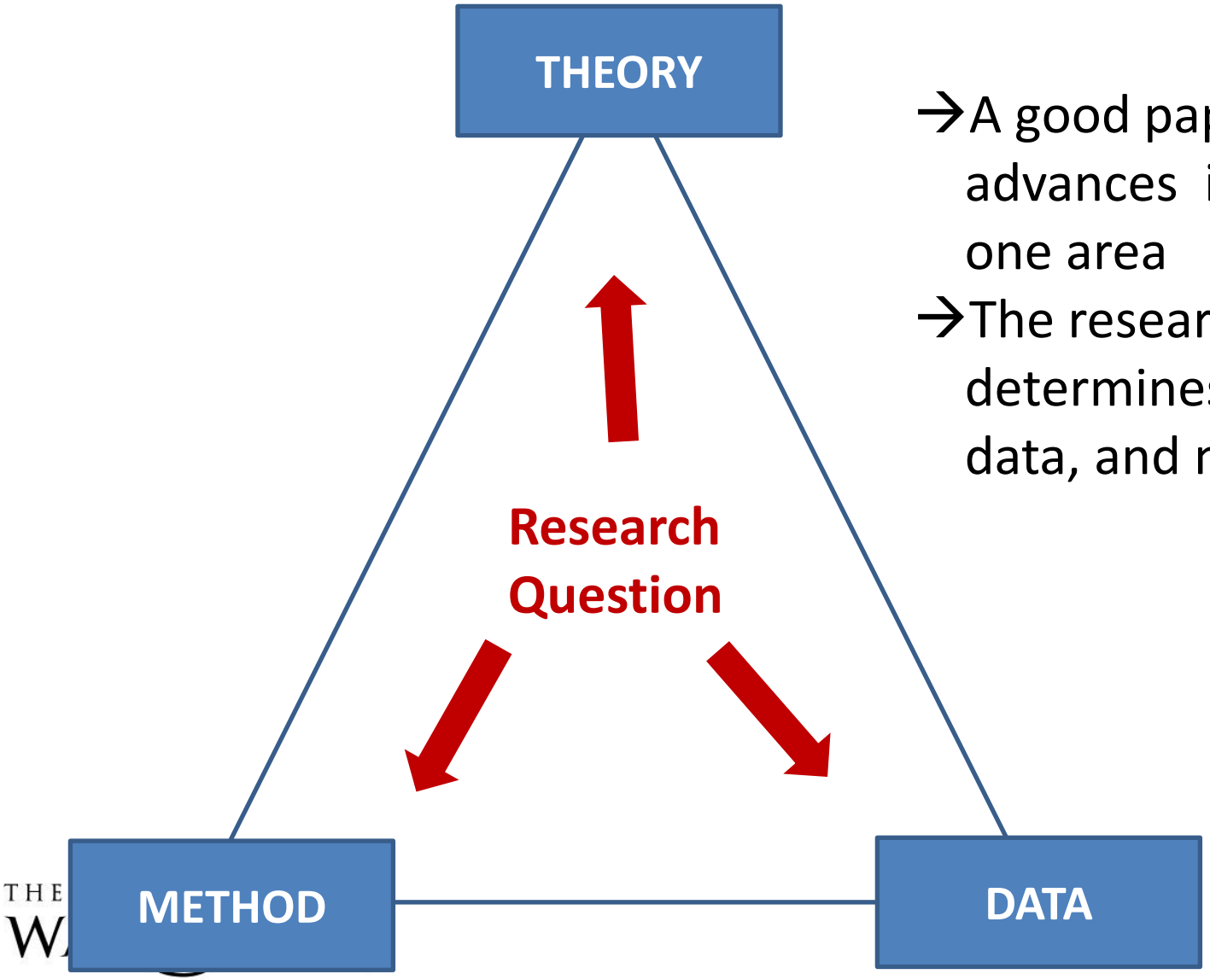


The Parable of the Elephant …

Six blind men go to observe an elephant. One feels the side and thinks the elephant is like a wall. One feels the tusk and thinks the elephant is a like a spear. One touches the squirming trunk and thinks the elephant is like a snake. One feels the knee and thinks the elephant is like a tree. One touches the ear, and thinks the elephant is like a fan. One grasps the tail and thinks it is like a rope. They argue long and loud and though each was partly in the right, all were in the wrong.



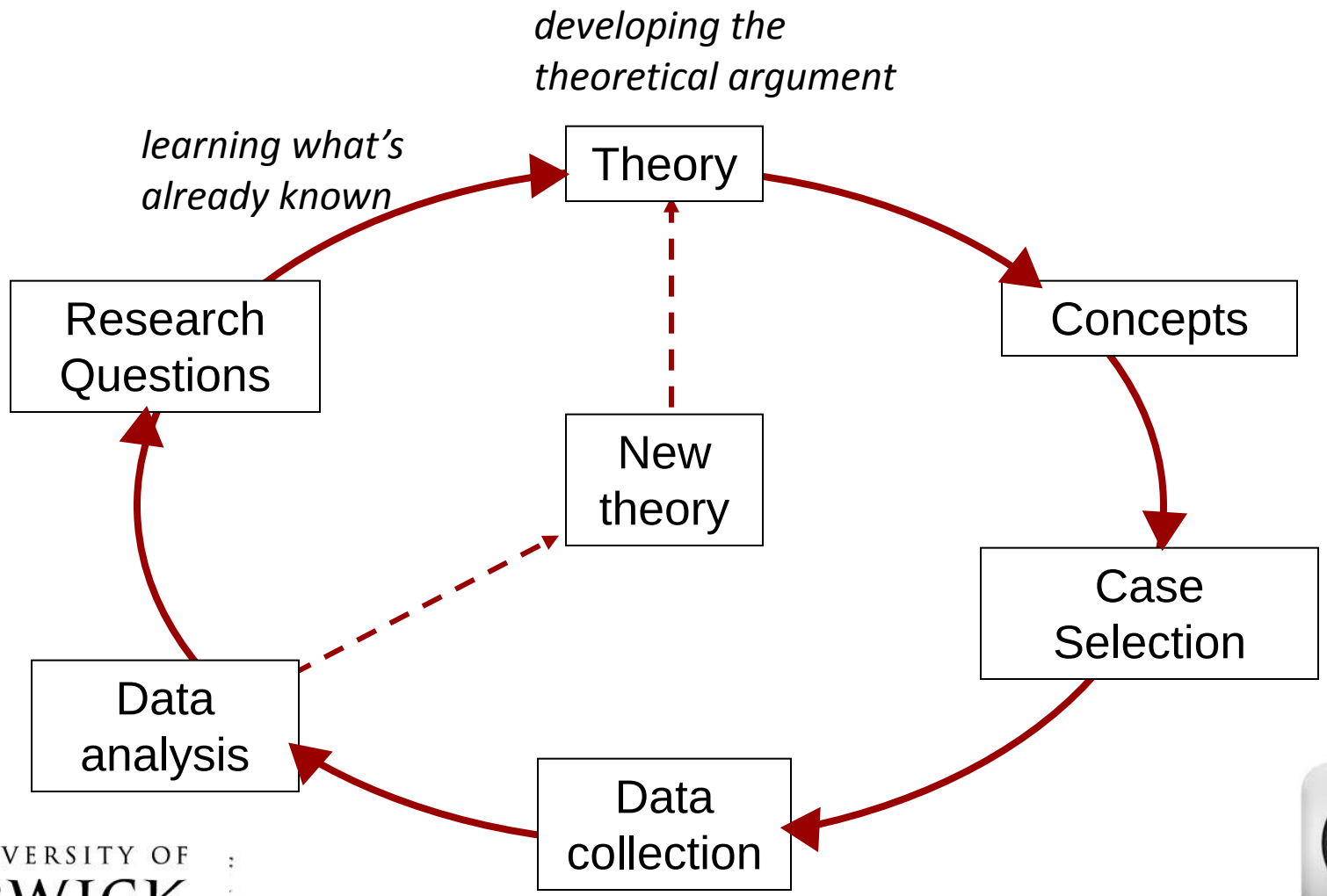
Triangle of scientific work



→A good paper advances in at least one area

→The research question determines theory, data, and method

Designing political science research



THE UNIVERSITY OF WARWICK

Your project

Attitudes toward ….

- UK EU-membership
- Democratic rules and procedures
- Capitalism
- Trade unions
- Political Islam
- Development aid
- Military interventions
- Women wearing headscarves
- …

Exercise for today

1) Explain, as best you can, the question you would like to explore. Include relevant context (< 250 words)

2) Based on the methodological readings for week 2 (but also week 1), explain how the implications of Geddes (1990) regarding selection on the dependent variable can (or cannot) be reconciled (< 1000 words)

Bowling for Columbine & Comparative Method

Question: *Why are there so many murders in America and not in the rest of the world?*

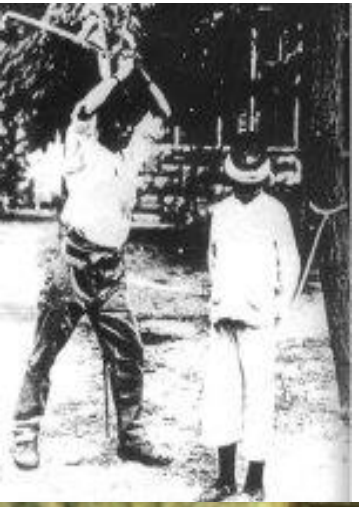
- What is the **basic argument** being put forth by Moore? That is, what is his **thesis**?
- What sort of **comparisons** does Moore use to support his thesis?
- Are the comparisons **useful**? how so?
- What type of **evidence** underlies his comparisons? Is the evidence sufficient, reliable, and valid?

THE UNIVERSITY OF WARWICK

Testing arguments about gun violence

Possible causes of America's high rate of gun violence

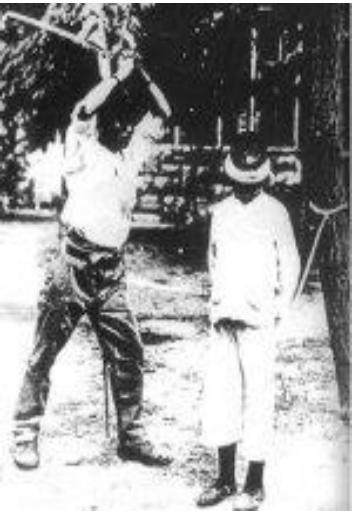
- Violent video games and entertainment
- Anti-social rock music and/or Marilyn Manson
- Too many guns
- Too much poverty
- Too much ethnic/racial diversity
- History of violence in the United States



Testing arguments about gun violence

Possible causes of America's high rate of gun violence

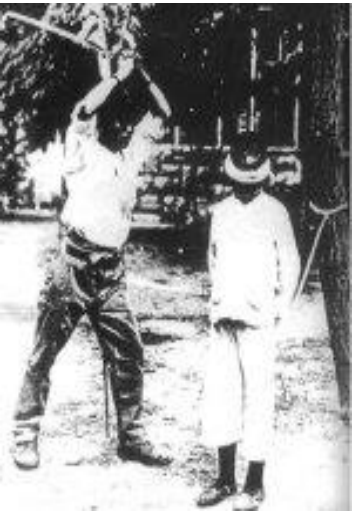
- ~~Violent video games and entertainment~~
- Anti-social rock music and/or Marilyn Manson
- Too many guns
- Too much poverty
- Too much ethnic/racial diversity
- History of violence in the United States



Testing arguments about gun violence

Possible causes of America's high rate of gun violence

- ~~Violent video games and entertainment~~
- ~~Anti-social rock music and/or Marilyn Manson~~
- Too many guns
- Too much poverty
- Too much ethnic/racial diversity
- History of violence in the United States

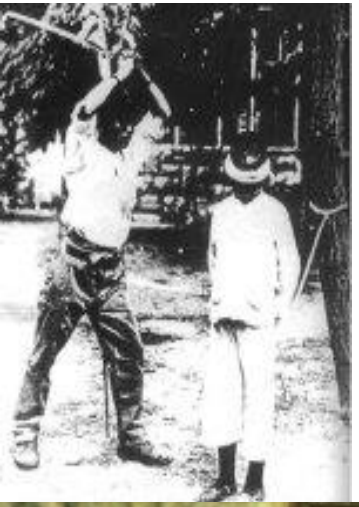


Testing arguments about gun violence

Possible causes of America's high rate of gun violence

- ~~Violent video games and entertainment~~
- ~~Anti-social rock music and/or Marilyn Manson~~
- ~~Too many guns~~
- …

→Through comparative checking, many possible **causes** of gun violence can be *eliminated* or problematized



Control through comparative method

Focus on similarities and differences to test an argument, i.e. to check whether claims about certain phenomena are valid

→ **“comparative checking”**

… to see if a variable of interest has a *similar effect* across a range of cases

WHY do we compare?

general purpose			
	comparing to control	**comparing to understand**	**comparing to explain**
basic strategy or purpose	comparative checking	interpretation	analytical Induction
logic or approach to comparative analysis	researcher uses a range of cases as a way to "test" a specific claim, hypothesis, or theory.	researcher is primarily interested in a single case and uses different cases or general theories as a way to learn more about the case he/she is studying.	researcher uses cases as a way to build a stronger theoretical explanation. cases are used in a "step-by-step" manner, with each case contributing to the development of a general theory.

Timothy Lim (2010): *Doing Comparative Politics*

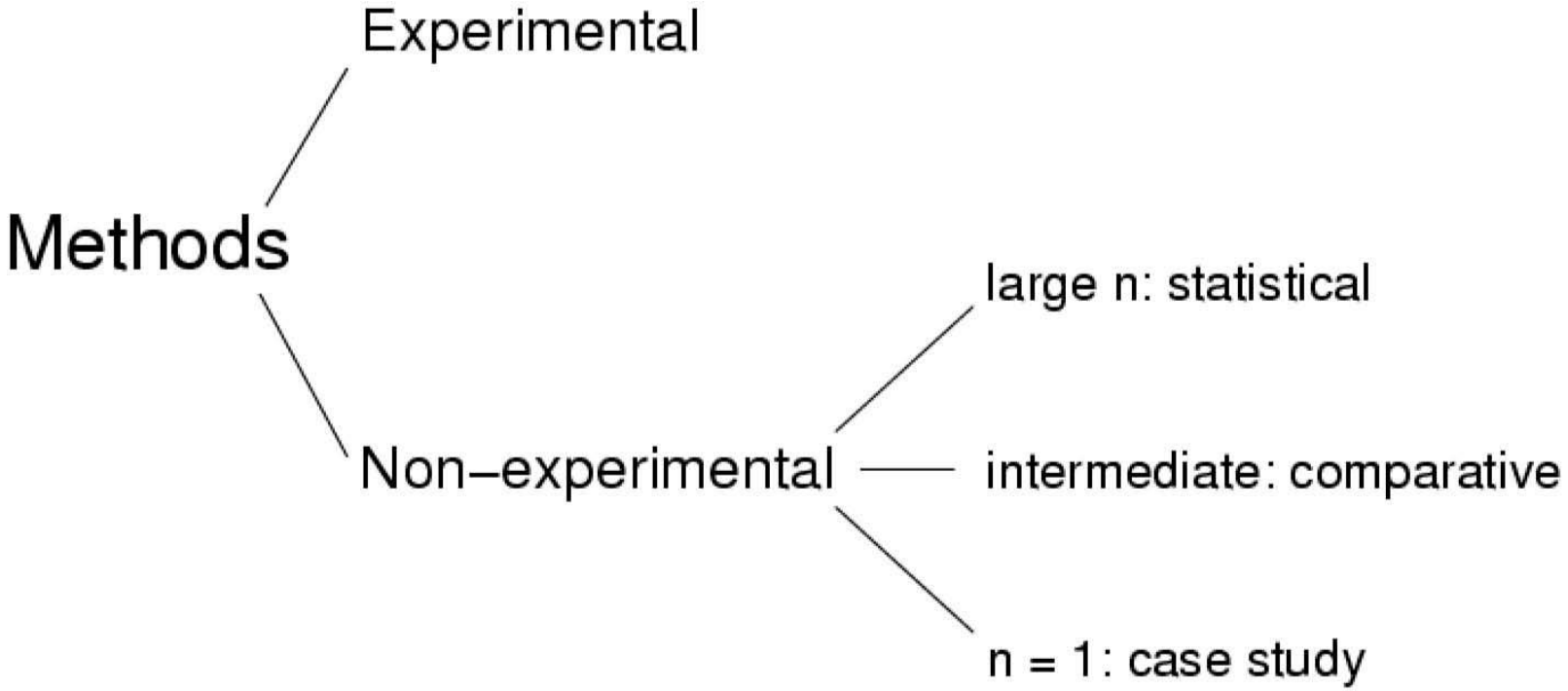
Political Science as *Science*

How Do You Know If You're Right?

Short answer: We have to be able to "test" the argument in some manner

- In the natural sciences, this testing is often (though not always) done through **experimentation** (or the **experimental method**), that is, the creation of carefully controlled conditions in which certain variables can controlled for in order to isolate others
- In the social sciences, "testing" is often done indirectly through **comparative analysis** or the **comparative method**

Comparative Methodology



Comparative approaches

high

Level of abstraction

low

3. Many units of analysis

2. A few units of analysis

1. Single unit case

one

scope

many

Single Case Studies

The study of a single case is considered comparative if it uses or develops concepts applicable to other cases, and/or seeks to make larger inferences.

Ideal to examine “deviant cases,” to generate hypotheses, to develop new classifications.

Inferences based upon one case are less secure.

How to compare?

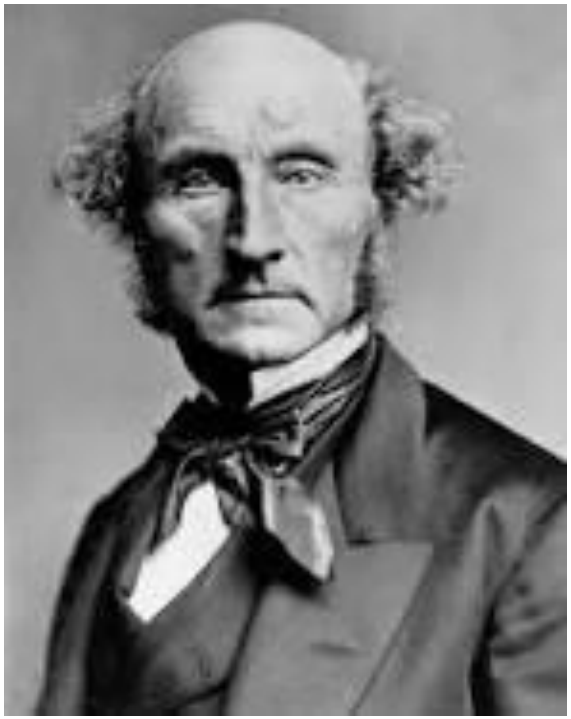
Selection on the

- Independent variables
- Dependent variable

Comparative case study design

The Method of Difference (MSS)

The Method of Agreement (MDS)

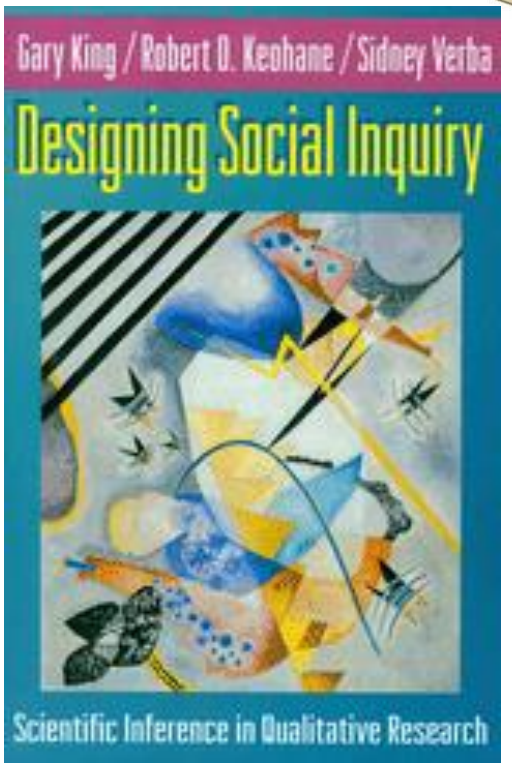


1843, *A System of Logic*

Selecting Cases on the Independent Variable

"selecting your cases according to the values of the independent variable that they take on."

- you have to know a little bit about all of your potential cases.
- you cannot act as if you also know the values that the dependent variable takes on.



KKV:
King, Keohane and Verba

Most Similar Systems (MSS) design

The Method of Difference

If, within the systems we are comparing, there is an occurrence and a non-occurrence of the phenomenon, and the circumstances in which these are observed are the same in all factors save one, then that one is the cause of the occurrence

Selection of cases that take on similar values of confounding variables, but different values of a key independent variable.

Confounds are "holds constant" because they take on the same values in all of the cases.

This is the design recommended by King, Keohane, and Verba.

Most Different Systems (MSD) Design

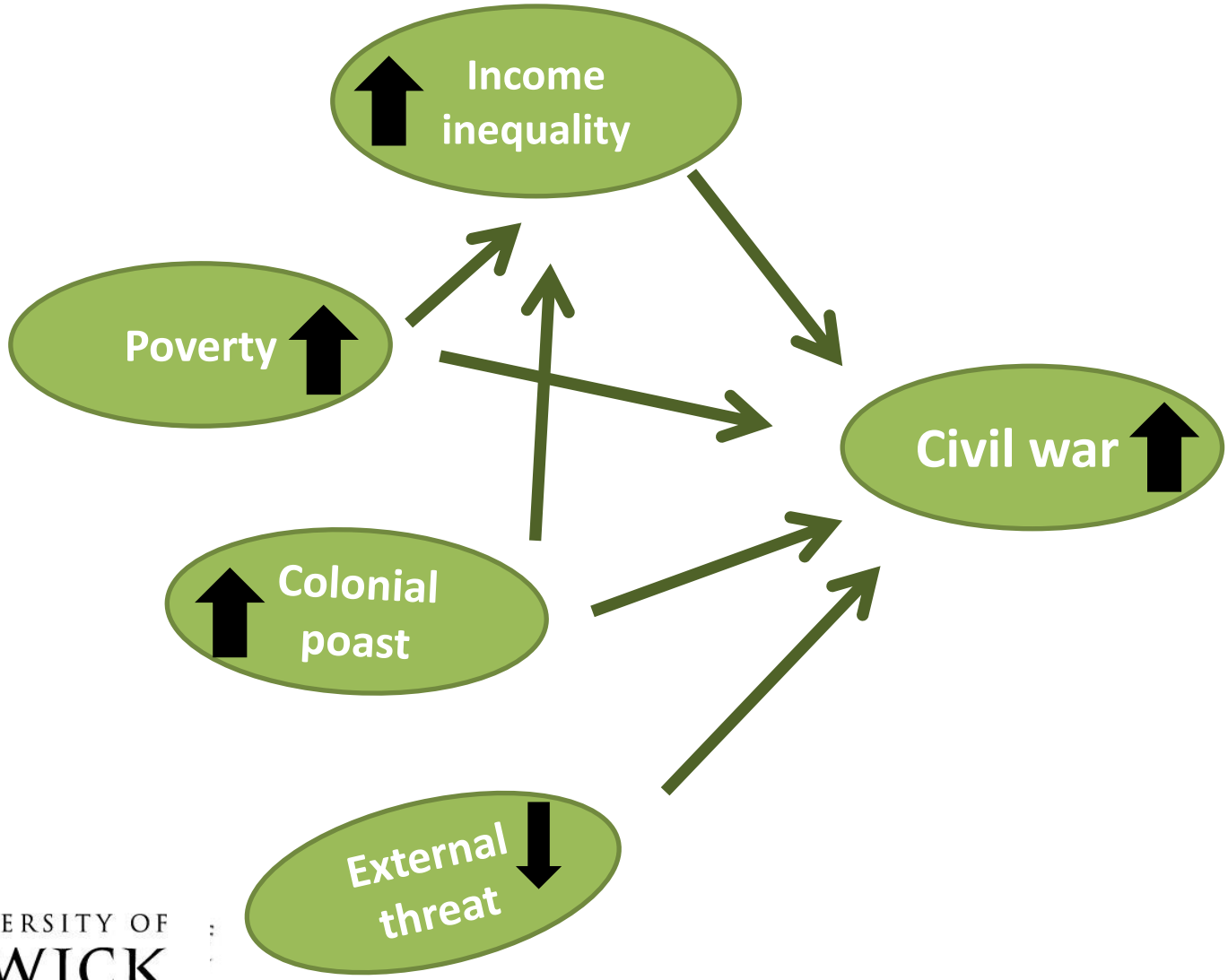
The Method of Agreement

If, within the systems we are comparing, the phenomenon we are interested in explaining have only one of several possible causal circumstances in common, then the circumstance in which all the instances agree is the cause of the phenomenon.

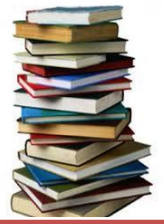
Selection of cases that take on very different values for multiple independent variables. If it turns out that these cases all take on the same value of a dependent variable, then we can rule out the independent variables as causes of the dependent variable.

<u>Less useful since it can only disprove a hypothesis.</u>

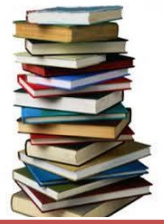
Example: Income inequality and civil war



Case	IV1: Income inequality	IV2: Poverty	IV3: Colonial past	IV4: External threat
Costa Rica	Moderate	Yes	Yes	No
El Salvador	High	Yes	yes	No
Cuba	high	yes	Yes	No



Case	IV1: Income inequality	IV2: Poverty	IV3: Colonial past	IV4: External threat	DV: Civil war?
Costa Rica	Moderate	Yes	Yes	No	No
El Salvador	High	Yes	yes	No	Yes
Cuba	high	yes	Yes	No	yes



Case	IV1: Income inequality	IV2: Poverty	IV3: Colonial past	IV4: External threat	DV: Civil war?
Costa Rica	Moderate	Yes	Yes	No	No
El Salvador	High	Yes	yes	No	Yes
Cuba	high	yes	Yes	No	yes

We can hold the confounds constant by selecting similar cases, here from Latin America.

It appears that income inequality does lead to civil war.

Selecting Cases on the Dependent Variable

Selecting cases according to the value of the dependent variable that they take on is more controversial than selecting on the independent variable.

It allows you to look at extreme values or *divergent cases*.

"However, if this design is to lead to meaningful … causal inferences, it is crucial to select observations without regard to values of the explanatory variables. K.K.V."

Method of Agreement

Selection of cases with same values of the dependent variable.

- This helps you to rule out possible causes, because independent variables that vary over these cases can't cause the dependent var.
- This method can only disprove a hypothesis, because it can't find a correlation

Case	Early industrialization?	Viable Socialist Party?
France	No	Yes
Britain	Yes	Yes

We can rule out "early industrialization" as a cause of whether a country has a viable socialist party.

Method of Difference

Selection of cases that take on different values of the dependent var.

After you have selected your cases, you determine what values they take on for some independent variables. Perhaps one independent variable will vary across your cases, and explain the D.V.

Case	Early industrialization?	Feudalism?	Viable Socialist Party?
France	No	Yes	Yes
Britain	Yes	Yes	Yes
USA	Yes	No	No

THE UNIVERSITY OF WARWICK

Adding a country without viable socialist party can add causal leverage to our earlier investigation.

	MSSD				MDSD		
	C1	C2	C3		C1	C2	C3
Features	a	a	a		a	d	g
	b	b	b		b	e	h
	c	c	c		c	f	i
Key explanatory factor	x	x	Not x		x	x	x
Outcome to be explained	y	y	Not y		y	y	y

THE UNIVERSITY OF WARWICK

Multiple Causation

Case	Independent variables	Dependent variable
Case 1	Context A (a,b,c,D)	Outcome X
Case 2	Context A (a,B,c,D)	Outcome Y

It is not always – or indeed often – the case that *one* factor alone is responsible for causing a phenomenon to occur. Mill's methods can obscure multiple causal factors.

From this table, what would we conclude is the causal factor?

Multiple Causation

Case	Independent variables	Dependent variable
Case 1	Context A (a,**b**,c,D)	Outcome X
Case 2	Context A (a,**B**,c,D)	Outcome Y

It is not always – or indeed often – the case that *one* factor alone is responsible for causing a phenomenon to occur. Mill's methods can obscure multiple causal factors.

But, couldn't it also be the case that it is the combination of A + B that is causing X?

→ The methods of difference and agreement can lead us to incorrect conclusions

Necessary and Sufficient Conditions

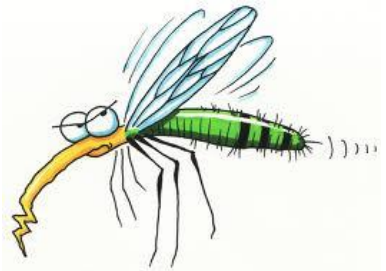
A **necessary** condition is a condition that must be present in order for some outcome to occur. But, its presence does not guarantee that the outcome will occur.

→ Oxygen is necessary to start a fire, but it is not sufficient by itself

A **sufficient** condition is a condition whose presence is sufficient for the phenomenon to occur. It is enough to get the job done, but it might not be necessary.

→ 'un-friending' your ex on facebook after you break up. 'Breaking up' is sufficient for you to un-friend him/her, but it isn't necessary. You could unfriend him/her for other reasons

Conditions can also be **both** necessary and sufficient



→ Being bitten by a mosquito carrying malaria is both necessary and sufficient for you to contract the disease

Methods of difference and agreement might not be able to identify sufficient conditions or to definitively establish a causal link.

THE UNIVERSITY OF WARWICK

Exercise for today

1) Explain, as best you can, the question you would like to explore. Include relevant context (< 250 words)

2) Based on the methodological readings for week 2 (but also week 1), explain how the implications of Geddes (1990) regarding selection on the dependent variable can (or cannot) be reconciled (< 1000 words)

The case selection bias debate

Geddes: Selecting cases based on the dependent variable biases conclusions. It can lead the researcher to perceive a causal relationship that doesn't exist

Collier and Mahoney: There is a problem with selecting on the dependent variable, but it is the opposite one: it can obscure causal relationships that actually exist

Geddes' Argument

If information is only collected on cases that exhibit a specific outcome and not on those that don't, we cannot know whether the factors identified really vary with the outcome

It is possible that there is no relationship between the identified cause and the observed effect. So while we can identify plausible variables we cannot test the theories.

BUT – others have responded and noted that this problem only applies to studies that are looking for sufficient conditions. If we are looking for necessary conditions, then this approach is entirely appropriate.

Example: Skocpol, *States and Social Revolution*

The argument:

State crisis (independent variable) → social revolution (dependent var.).

The Criticism:

By only selecting cases that experienced social revolution, she misses the fact that there are many other cases that have experienced state crisis but not experienced social revolution. She exaggerates the relationship between state crisis and social revolution.

	Revolution	**No Revolution**
Defeated or Lost Territory	Bolivia Defeated 1935 Revolution 1952	Peru, 1839 Bolivia, 1839 Mexico, 1848 Paraguay, 1869 Peru, 1883
Not Defeated within 20 Years	Mexico, 1910 Nicaragua, 1979	All Others

… matter of perspective

If Skocpol is claiming that state crisis is *sufficient* to cause social revolution, her study suffers from selection bias

BUT – if she is claiming that state crisis is simply a *necessary* condition, then her research design still holds up.

	Revolution	**No Revolution**
Defeated or Lost Territory	Bolivia Defeated 1935 Revolution 1952	Peru, 1839 Bolivia, 1839 Mexico, 1848 Paraguay, 1869 Peru, 1883
Not Defeated within 20 Years	Mexico, 1910 Nicaragua, 1979	All Others

Case selection affects types of inferences

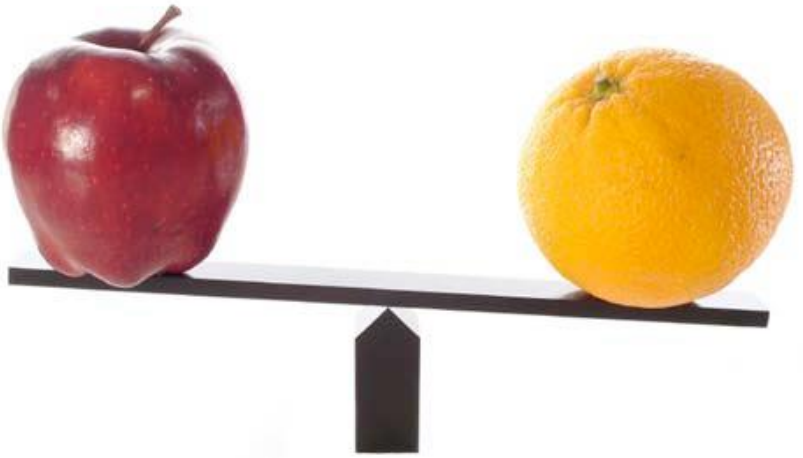
- State crisis is a necessary condition for social revolution
  - A good design to test this claim would be to compare countries that have experienced social revolution and see if they all experienced state crisis (selecting on the dependent variable)

- State crisis is a necessary and sufficient condition for social revolution
  - A good design here would be to look for state crises and see if they all lead to social revolution (selecting on the independent variable)

WHAT do we compare?

Can we compare **apples** and **oranges**?

(why do many people consider apples and oranges non-comparable?)



WHAT do we compare?

… if apples and oranges can be compared can, say, **Haiti** and the **United States** also be compared? Why or why not?

WHAT is comparable?

Key point: There is no fixed answer.

(1) The answer always **depends** on the research question

(2) Comparisons focus on internal or domestic political structures, actors, and processes including

- **events** (e.g., wars or revolutions)
- **political or social institutions** (e.g., the executive branch, the military, economic agencies)
- **policies** (e.g., health care, educational policies, welfare)

(3) We can compare "entities whose **attributes** are in part shared (similar) and in part non-shared"; dimensions are multiple: spatial, temporal, and functional.

Roadmap

Part I - **RESEARCH DESIGN**

- Causal thinking and research designs
- Comparative analysis and case selection

Part II - **DATA COLLECTION**

- **(Semi-)Structured Interviews**
- Observational research and ethical questions
- Observational research and ethical questions

Part III - **DATA ANALYSIS AND CAUSAL INFERENCE**

- Case studies and process-tracing
- Qualitative Comparative Analysis (QCA)
- The mixed-methods approach

Exercise for next week

1) On your own or with one or two fellow students, prepare an interview protocol to investigate your specific research question, and conduct at least two semi-structured interviews each using this protocol. *Keep your interview notes!* Place in the dropbox your research question and interview protocol. If this exercise is done as teamwork, I encourage you to organize one observed interview each, so that a fellow student can provide feedback on interview style.
2) Go to one top political science journal (e.g. American Journal of Political Science, International Organization, Comparative Political Studies) and select three articles in the most recent issue. Based on the abstract only, what are the cases in the presented study? Think about the external and internal validity of the respected study.

Departmental seminar workshops

http://www2.warwick.ac.uk/fac/soc/pais/research/paisseminars/

Tomorrow. Pempel: 'The Economic-Security Nexus in Northeast Asia'
[5-6:30 pm in S0.18, introducted by Chris Hughes]

Methodologically relevant workshops (save the date!)
22/10. Julian Wucherpfennig – reversed causality & instrumental variables
02/12. Jale Tosun – operationalization of dependent variables
10/12. Matthew Wilson – Sequence analysis
21/01. Jan-Hinrik Meyer-Sahling – index building
04/02. Kristian Gleditsch
06/05. Bernd Schlipphak
18/05. Paul Heywood

THE UNIVERSITY OF WARWICK

THE UNIVERSITY OF WARWICK

3: (Semi-)Structured Interviews

Tina FREYBURG

[Introduction to Qualitative Methods]