POLS2044 WEEK 4
Concepts and Measures

Australian National University
School of Politics & International Relations
Dr. Richard Frank

Last week we discussed different research design approaches. In Week 4 we grapple with the opportunities and difficulties when trying to connect our theoretical concepts to the real world that we can measure (whether with words or numbers).

My goals for Week 4:

1. Highlight important considerations when moving from theory to evidence
2. Have you grapple with the challenges and opportunities involved in this process
3. Understand what is involved when we talk about measurement validity
4. Examine a few examples that you would think would be unproblematic
5. Apply what you learned about measurement to a particular topic in workshop

---

I. Reading

There is one reading for this week—Chapter 5: Measuring Concepts of Interest from the textbook.

This week is all about figuring out how to connect our causal theories to measured real-world evidence that has demonstrated reliability and validity while minimizing systematic bias. The chapter is relatively short and straightforward, but the concepts and research challenges it introduces are some of the most important and intractable ones in Political Science.

LECTURE PART 1: INTRODUCTION

Why should we care about measurement?

> We can fall in love with stories we tell ourselves about the world; however, if these stories remain untested assumptions, we have no idea whether they are true or false.

Today's motivating questions

> How can we link solid causal theories to real-world evidence?
>
> How can we be sure this evidence has measurement validity?

Motivating puzzle

> Most people use real-world data without thinking about how they are generated and whether they capture what they think they do.

LECTURE PART 2: Moving from concept to measure

Moving from theory to test

What is the difference between a theory and a hypothesis?

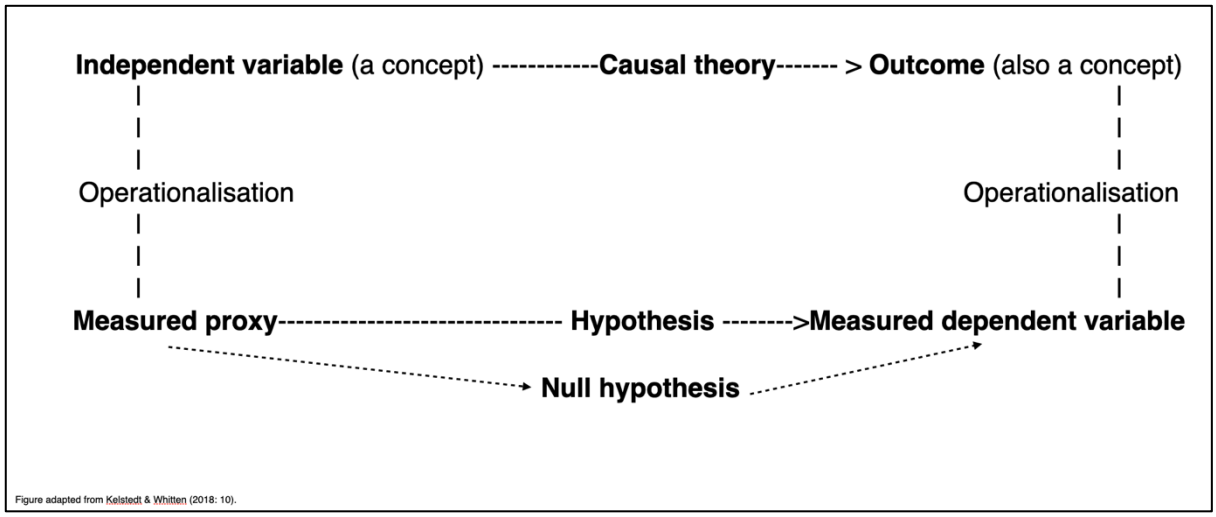


Figure adapted from Kelstedt & Whitten (2018: 10).

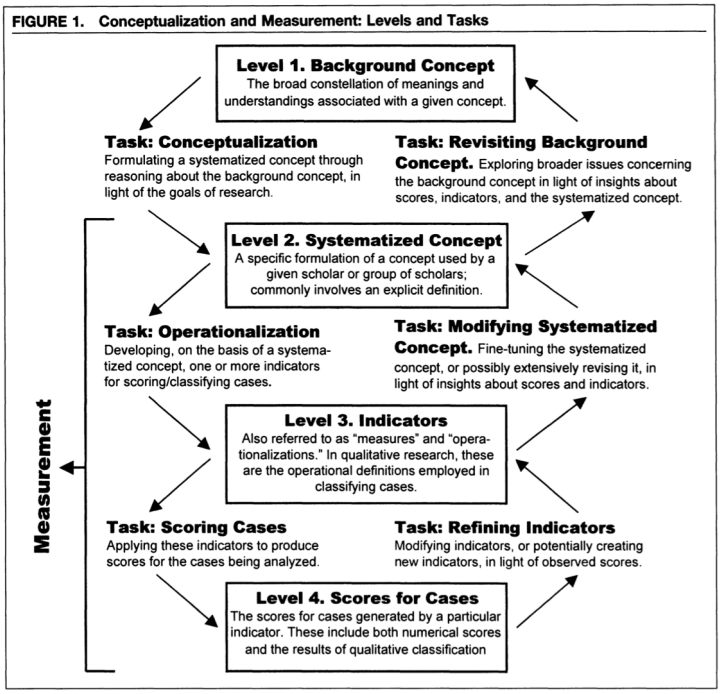
The null hypothesis is akin to the legal principle of the presumption of innocence.

Slide of tweet saying how the lack of data is making it hard to build early warning alerts for climate change.

From theory to measurement

Causal theories are relationships between concepts.
Measurement is trying to observe these concepts (or their proxies) in the real world.

Adcock and Collier (2001: 531) path diagram



In other words:

Step 1: Conceptualise
Step 2: Operationalise
Step 3: Code

Where is Canberra's best coffee shop?

1: Conceptualise coffee shop characteristics
2: Operationalise these characteristics
3: Code as many coffee shops as possible

Conceptual/coding challenges

Are observations heterogeneous or homogenous?

Short excerpt from This is Spinal Tap (1984)

https://youtu.be/KOO5S4vxi0o

Any gap between theory and measurement here?

Causal threat: measurement error

Measurement error is the difference between the measured value and the true value of something.

It includes both a random component and the potential of a systematic component

Type 1 and Type 2 errors

False positives and false negatives

Random error

Erik Gartzke. 1999. "War Is in the Error Term." *International Organization* 53(3): 567-587.

Random measurement error

"People are not very good at understanding randomness. There's much more chance out there than we think there is. While we are seeking for patterns and explanations as we look backward, we're not giving a fair shot to the explanation that many things are really just random events."
— Lisa Goldberg

Pareidolia—Seeing things that are not there.

New Hampshire's Old Man of the Mountain

Challenges to measurement

Conceptual clarity—Do we know what we want to measure?

Operational reliability—Are the measures repeatable and consistent?

Conceptual validity—Does the measure accurately measure the concept we are trying to measure?

Validity

Face validity

On its face does a measure appear to be measuring what it says it is measuring?

Example of North Korea's constitution

Content validity

Does a measure capture all of the systemised concept? Is anything missing? Is anything there that should not be?

Example: Polity V does not include measures of participation

Criterion validity

Does a measure correlate with criterion (i.e., ground truth) variables?

Examples of "Le Grand K" and political polling

Construct validity

Do measures behave the way you theoretically expect in the wild?

Example of the relationship between governance and democracy

Validity and reliability

Can you think of a valid but unreliable measure?
A reliable but invalid measure?

LECTURE PART 3: A few examples

Example 1: GDP

Definition from the World Bank

(https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?view=chart)

CNBC video about the GDP definition

(https://youtu.be/iLom1WlqwS0)

Joseph Stiglitz article in Scientific American "GDP is the wrong tool for measuring what matters"

Bhutan's Gross National Happiness Index

(https://ophi.org.uk/policy/gross-national-happiness-index/#:~:text=The%20phrase%20'gross%20national%20happiness,approach%20towards%20notions%20of%20progress)

Population Definition from the World Bank

(https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?view=chart)

Australia's definition of population

(https://population.gov.au/population-topics/topic-population-measurement)

Article example of looking at Sweden's population according to a subnational grid

Example of PRIO-GRID 2.0's grid population data

Example: Democracy

When is a country a democracy?

How is democracy a latent or unobservable concept?

V-DEM's conceptual scheme

Munck & Verkuilen's (2002: 8) process

Table 2
A Framework for the Analysis of Data: Conceptualization, Measurement, and Aggregation

Challenge	Task	Standard of Assessment
Conceptualization	Identification of attributes	Concept specification: Avoid maximalist definitions (the inclusion of theoretically irrelevant attributes) or minimalist definitions (the exclusion of theoretically relevant attributes)
	Vertical organization of attributes by level of abstraction	Conceptual logic: Isolate the "leaves" of the concept tree and avoid the problems of redundancy and conflation
Measurement	Selection of indicators	Validity: Use multiple indicators and establish the cross-system equivalence of these indicators; use indicators that minimize measurement error and can be crosschecked through multiple sources
		Reliability
	Selection of measurement level	Validity: Maximize homogeneity within measurement classes with the minimum number of necessary distinctions
		Reliability
	Recording and publicizing of coding rules, coding process, and disaggregate data	Replicability
Aggregation	Selection of level of aggregation	Validity: Balance the goal of parsimony with the concern with underlying dimensionality and differentiation
	Selection of aggregation rule	Validity: Ensure the correspondence between the theory of the relationship between attributes and the selected rule of aggregation
		Robustness of aggregate data
	Recording and publicizing of aggregation rules and aggregate data	Replicability

Note focus on:
1. Conceptualisation
2. Measurement (including reliability, validity, and replicability)
3. Aggregation (often overlooked)

Munck & Verkuilen's (2002: 13) process

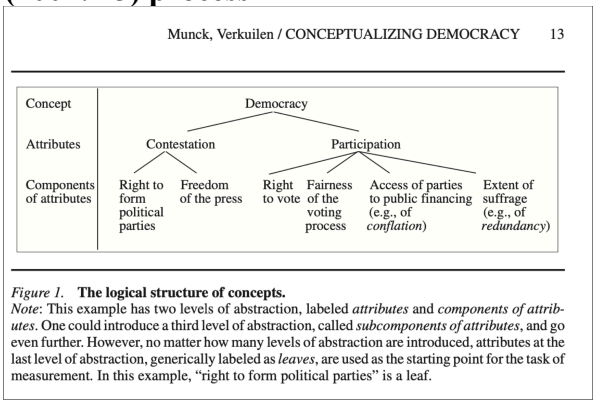


Figure 1. **The logical structure of concepts.**
Note: This example has two levels of abstraction, labeled *attributes* and *components of attributes*. One could introduce a third level of abstraction, called *subcomponents of attributes*, and go even further. However, no matter how many levels of abstraction are introduced, attributes at the last level of abstraction, generically labeled as *leaves*, are used as the starting point for the task of measurement. In this example, "right to form political parties" is a leaf.

Example 4: Natural resources and civil conflict onset

Taken from Humphreys (2005)

Mechanisms connecting natural resources & civil conflict onset

Mechanisms

1. Greedy rebels
2. Greedy outsiders
3. Grievances
4. Feasibility
5. Weak states
6. Sparse networks

Proxies

1. Diamond production
2. State instability
3. State instability x autocracy
4. Oil reserves (per capita)
5. Oil production (per capita)
6. share of agriculture (% GDP)

Measurement and transparency

Is reporting data a real measure of governmental transparency? Or capacity?

Today's motivating questions

How can we link solid causal theories to real-world evidence?

How can we be sure this evidence has measurement validity?

III. WEEK 4 WORKSHOP

The focus of today's workshop is on applying the readings and lecture material.

Item 1: Individual work

The first part of class is trying to get us to think about an underlying concept and how we might measure it in the world. For this week, the concept is: **corruption**. This is a tough one to measure because those that are corrupt have every incentive to hide their corruption from anyone who could punish them for it either legally or reputationally.

1. **Write five questions or prompts that you could use when interviewing subjects about their experience with governmental corruption.**

Submit your questions to Wattle/POLS2044/Week 4/Item 4.1 before you move on to the group work.

Item 2: Small group work

Next come together in groups of three/four students. Read off your questions/prompts to each other.

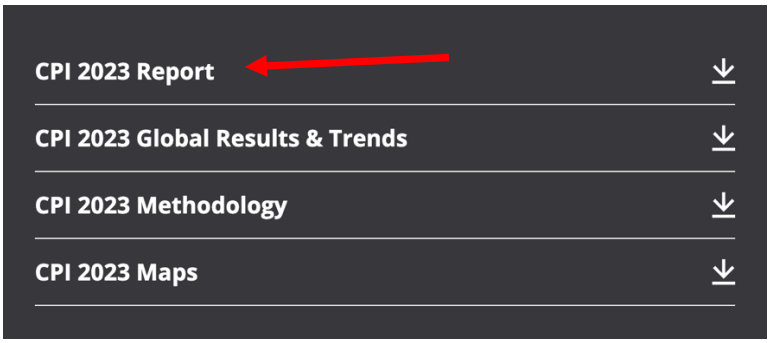
2. **Are there any notable similarities or difference between your responses?**

3. **How did your questions map onto your underlying assumptions about the relevant concepts and possible measurements of "corruption"?**

Submit your questions to Wattle/POLS2044/Week 4/Item 4.2 before you move on. Be sure to have one student submit for the group and include the other student's names & uniid.

Item 3: Small group data activity

Now go to Transparency International's (TI) most recent report on global corruption (https://www.transparency.org/en/cpi/2023). The full report download link is towards the bottom of the page.

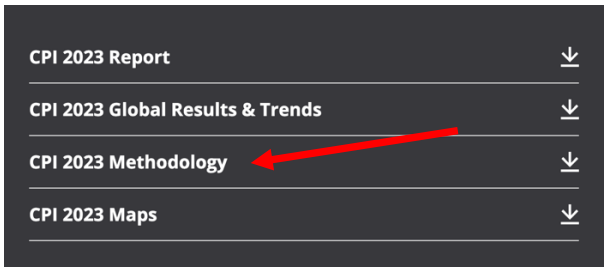


Look at some of the country rankings and then the methodology section (p. 20).

4. **Is this measurement methodology clear?**

5. **Do you think it is methodologically valid and reliable?**

Maybe we are being unfair or are missing the underlying links between concept and measurement. If you download the "CPI 2023 Methodology" PDF, you can find and extract their "Technical Methodology Note" which has more detail.

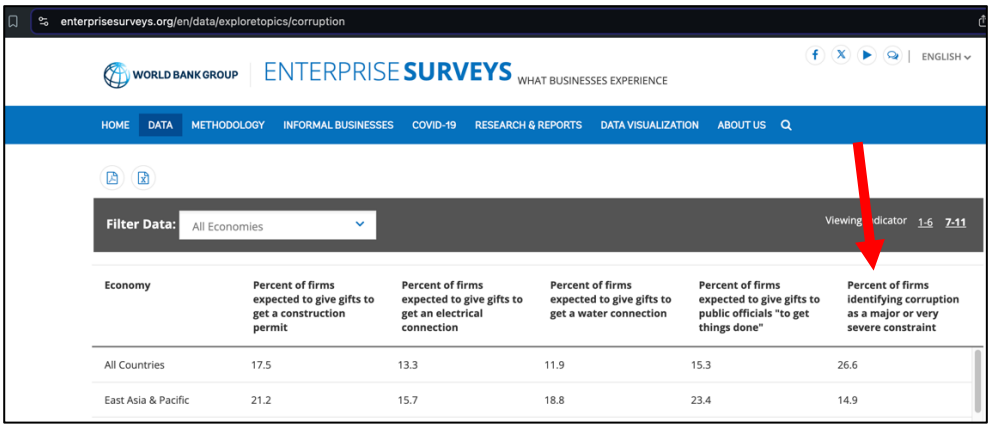


6. **Does this level of detail help clarify things?**

7. **What do you think about Transparency International's aggregate efforts?**

Finally, go to the World Bank's Enterprise Surveys website which focuses on their corruption surveys (https://www.enterprisesurveys.org/en/data/exploretopics/corruption).

Try sorting the summary table by indicator 11 ("Percent of firms identifying corruption as a major constraint").



8. **Are there any notable similarities or differences between this ranking and that of TI?**

9. **More generally, do you think these questions link more (or less) directly to your conceptualisation of corruption in your questions you started with for question 1 above?**

Submit your questions to Wattle/POLS2044/Week 4/Item 4.3 before you move on. Be sure to have one student submit for the group and include the other student's names & uniid.

Finally, as you wrap up, a few questions to leave you thinking about (no need to submit responses on Wattle). What were the main takeaways of your individual and small group activities conceptualising, measuring, and aggregating corruption? What connections can you draw between the lecture and workshop discussion of concepts and measures to your own research?