POLS 205
Political Science as a Social Science

Concepts & Measurement

Christopher Adolph

University of Washington, Seattle

April 21, 2010

Outline

From Concept to Measure

Valid & Reliable Measurements

Levels of Measurement

From concept to measure

Political scientists distill their observations about the world into theories about concepts.

Vague observation:

Many countries ruled by dictators seems to have many poor citizens, quite a few rich citizens, but relatively few middle class citizens.

Refined to show concepts & causality:

Democracy reduces economic inequality.

From concept to measure

Democracy reduces economic inequality.

But how do we define the concepts in our theories?

How do we make the concept *clear* and *sharp*, so that we know whether any evidence we gather supports the theory?

What is democracy?

Most popular definition (Dahl):

Democracy is the regular *contestation* for political leadership through the *participation* of the largest possible share of the polity's members.

From concept to measure

Democracy is the regular *contestation* for political leadership through the *participation* of the largest possible share of the polity's members.

Note two crucial concepts:
Contestation: The degree to which political power is "up-for-grabs", e.g., through free, fair, and open elections.

Participation: The fraction of the public with a meaningful say in the contestation of power

From concept to measure

Which *elements* of the concepts are relevant for the current theory?

We might recall the median voter theorem, and speculate that

1. The voters (or selectorate, in non-democracies) choose the leader
2. The degree of redistribution is *the* major issue dividing the voters; the poorer the voter, the more redistribution desired
3. By the MVT, the poorer the *median* voter, the more redistribution will happen
4. As we expand the electorate/selectorate, the wealth of the median voter declines

From concept to measure

This yields our refined hypothesis:

Countries with more encompassing voter franchises will have less post-tax and transfer economic inequality.

To test our theory, we need to not only measure democracy, we need to isolate and measure *only* the relevant parts of the concept of democracy:

the rate of *participation* in societies with high levels of *contestation*

Example: Polity IV

Example: Polity IV Measure of Democracy

From Polity IV brief summary

The Polity conceptual scheme is unique in that it examines concomitant qualities of democratic and autocratic authority in governing institutions, rather than discreet and mutually exclusive forms of governance. This perspective envisions a spectrum of governing authority that spans from fully institutionalized autocracies through mixed, or incoherent, authority regimes (termed "anocracies") to fully institutionalized democracies. The "Polity Score" captures this regime authority spectrum on a 21-point scale ranging from -10 (hereditary monarchy) to +10 (consolidated democracy). The Polity scores can also be converted to regime categories: we recommend a three-part categorization of "autocracies" (-10 to -6), "anocracies" (-5 to +5 and the three special values: -66, -77, and -88), and "democracies" (+6 to +10)

`www.systemicpeace.org/polity/polity4.htm`

Example: Polity IV Measure of Democracy

According to the Polity IV measures (Ted Gurr et al), democracy consists of four components:

1. Regulation of executive recruitment
2. Competitiveness of executive recruitment
3. Openness of executive recruitment
4. Constraints on chief executive

Example: Polity IV Measure of Democracy

Each can be scored on a scale, or a set of *levels* of measure.
E.g., *Regulation of executive recruitment* is scaled:

- +3 regular competition between recognized groups
- +2 transitional competition
- +1 factional or restricted competition
- 0 no competition

While constraints on the executive are scaled:

- +4 Executive parity or subordination
- +2 Substantial limitations
- 0 No limitations

What do we notice about this scaling?

Two criteria for good measurement

The measurement of a social concept should be valid and reliable.

Validity The measurement faithfully repoduces and isolates the concept required by the hypothesis

Reliabilitity The measurement can be systematically gathered with low error by any competent scientist

Measurement validity

Three components of validity, with example theory: Ethnic division $\Rightarrow$ Lower welfare state spending:

Face validity On first inspection, the measurement seems to recapitulate the concept

Measurement validity

Three components of validity, with example theory: Ethnic division $\Rightarrow$ Lower welfare state spending:

Face validity On first inspection, the measurement seems to recapitulate the concept

Example: Ethno-linguistic fractionalization using self identified ethnicity

Measurement validity

Three components of validity, with example theory: Ethnic division $\Rightarrow$ Lower welfare state spending:

Face validity On first inspection, the measurement seems to recapitulate the concept

Example: Ethno-linguistic fractionalization using self identified ethnicity

Content validity The measurement attempts to capture all the relevant elements of the concept, and nothing extraneous

Measurement validity

Three components of validity, with example theory: Ethnic division $\Rightarrow$ Lower welfare state spending:

Face validity On first inspection, the measurement seems to recapitulate the concept

Example: Ethno-linguistic fractionalization using self identified ethnicity

Content validity The measurement attempts to capture all the relevant elements of the concept, and nothing extraneous

Example: Does ELF contain everything relevant? "Intensity" of division? Overlapping division? Identification by others?

Construct validity The measurement behaves like the concept would

Example: Does a substantial increase in ELF have the same effect on welfare as a substantial increase in ethnic division is *expected* to (note circularity)

Measurement validity trade-offs

Note some problems with trying to achieve all three kinds of validity:
For complex concepts, like “democracy” or “ethnic division”, high content validity means mushing together many different elements

This might be accomplished by an “index” where the units are non-comparable

In social science, an index is a single number calculated from a variety of variables. The units of this index may be either:

1. Clear and comparable: Dow Jones Industrial Average
2. Clear & non-comparable: Misery Index (Unemployment $+$ Inflation)
3. Vague and non-comparable: Polity IV

Measurement validity trade-offs

For a particular concept, content validity may seem to demand an omnibus (non-comparable) index, which threatens *construct* validity

Measurement validity trade-offs

For a particular concept, content validity may seem to demand an omnibus (non-comparable) index, which threatens *construct* validity

But we could always *break* the concept down until each component has its own variable

Measurement validity trade-offs

For a particular concept, content validity may seem to demand an omnibus (non-comparable) index, which threatens *construct* validity

But we could always *break* the concept down until each component has its own variable

Then all the content of the concept appears in some variable, but each variable might behave in a slightly different way, as expected for that component

Measurement validity trade-offs

For a particular concept, content validity may seem to demand an omnibus (non-comparable) index, which threatens *construct* validity

But we could always *break* the concept down until each component has its own variable

Then all the content of the concept appears in some variable, but each variable might behave in a slightly different way, as expected for that component

And we avoid trying to add "unemployment" and "inflation", or trying to decide how much we would need to expand voting rights to make up for a given reduction elected officials' power

(A good thing: trying to answer such questions may be *impossible'*, in the sense that no general answer exists)

How important is face validity?

Content and construct validity seem vital, though achieving both may require paring down our concepts

How important is face validity? Two questions:

1. Is face validity merely in the eye of the beholder (Justice Potter on obscenity: “I know it when I see it”)?

How important is face validity?

Content and construct validity seem vital, though achieving both may require paring down our concepts

How important is face validity? Two questions:

1. Is face validity merely in the eye of the beholder (Justice Potter on obscenity: “I know it when I see it”)?

2. Suppose we have content and construct validity. Do we still need face validity?

How important is face validity?

Content and construct validity seem vital, though achieving both may require paring down our concepts

How important is face validity? Two questions:

1. Is face validity merely in the eye of the beholder (Justice Potter on obscenity: "I know it when I see it")?

2. Suppose we have content and construct validity. Do we still need face validity?

   Example: Suppose my measure of potential seasonal migration from Mexico to California is the total rainfall in the agricultural areas of Mexico Is this facially valid?

How important is face validity?

Content and construct validity seem vital, though achieving both may require paring down our concepts

How important is face validity? Two questions:

1. Is face validity merely in the eye of the beholder (Justice Potter on obscenity: “I know it when I see it”)?

2. Suppose we have content and construct validity. Do we still need face validity?

   Example: Suppose my measure of potential seasonal migration from Mexico to California is the total rainfall in the agricultural areas of Mexico
   Is this facially valid?
   It is content & construct valid: low rainfall lowers crop yields which sends agricultural workers in search of work in California

Measurement reliability

Reliability: The measurement can be systematically gathered with low error by any competent scientist

The main problem for natural science: instrumentation error

In social science, our instruments are human

Reliability applied to:

A survey . . . reponses do not differ from one survey-taker to the next

An experiment . . . behavior does not depend on which staff administered the treatment

An observational study . . . behavior does not dependent on who collected & coded the (archival/interview) data

Measurement reliability

Inter-coder reliability: The percentage of items to different coders would code the *same*.

What could cause low inter-coder reliability?

1. Unskilled or unscrupulous coders
   Monitoring and cross-checking essential

2. Lack of *common* conceptual understanding
   Note this makes our codings *conditional* on a specific tacit understand, e.g., “the way political scientists think about democracy, or race, or education”

3. Ambiguous primary data
   Very hard to solve: choose a concept & measurement strategy that *avoids* this problem

Measurement challenges: from bad to worse

Easy? National economic production

Measurement challenges: from bad to worse

Easy?

National economic production
Unemployment

Measurement challenges: from bad to worse

Easy?	National economic production Unemployment
Not so easy?	Poverty

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain

Measurement challenges: from bad to worse

Easy?	National economic production Unemployment
Not so easy?	Poverty Physical pain
Hard	Political liberty

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy
	Congressional roll-call liberalism

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy
	Congressional roll-call liberalism
	Social capital

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy
	Congressional roll-call liberalism
	Social capital
Impossible?	Political “sophistication”

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy
	Congressional roll-call liberalism
	Social capital
Impossible?	Political “sophistication”
	Judicial “activism”

Measurement challenges: from bad to worse

Easy?	National economic production
	Unemployment
Not so easy?	Poverty
	Physical pain
Hard	Political liberty
	Democracy
	Congressional roll-call liberalism
	Social capital
Impossible?	Political “sophistication”
	Judicial “activism”
	Election “mandates”

Measurement challenges: from bad to worse

It's better to replace a loaded/controversial concept with a clear (but slightly different) one

Don't try to measure "sophistication" directly. Just measure *knowledge of uncontested political facts*.

(Note the difficulty in identifying such facts, and the danger than in a largely "unsophisticated" population, many important facts will be contested)

Don't measure "judicial restraint"; instead measure *Justices' probability of overturning ideologically opposed precedent*

Is there any concept we could replace *Electoral Mandates* with?

Continuous & discrete data

All data are either continuousor discrete

Determines which statistical tools are the right ones for your dependent variable

Discrete data can be matched up to the integers. There is a clear distinction between each possible value a discrete variable may take on.

Continuous data can take on any real value between a lower and upper bound, which might be $[-\infty, \infty]$.

Notice the symbol ∞, for infinity

Infinity is a tricky mathematical concept, but one tied up with the distinction between discrete and continuous variables

Aside: Integers, Real Numbers, & Infinity

Integers are the negative whole numbers, positive whole numbers, and zero:

$$-\infty, \ldots, -1000, -999, \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots, 999, 1000, \ldots, \infty$$

There are infinitely many integers

Aside: Integers, Real Numbers, & Infinity

Real numbers include every possible decimal within a given interval:

$$\mathbb{R} \in (\ell, u)$$

We can't list the real numbers, even using "..."

Between any two real numbers there are more real numbers.

In fact, there are an uncontable infinity of reals between any two reals!

Aside: Integers, Real Numbers, & Infinity

There are an uncontable infinity of real numbers between any two real numbers

Proven by Georg Cantor (19th c. mathematician)

Consider a sequence of numbers, like
$0.10000000, 0.02000000, 0.00300000, 0.00040000, 0.00005000, \ldots$

Cantor realize these numbers could be put in a table like so. . .

Aside: Integers, Real Numbers, & Infinity

Cantor's Diagonal Proof that Reals are Uncountably Infinite

0	.	1	0	0	0	0	0	0	...		1
0	.	0	2	0	0	0	0	0	...		2
0	.	0	0	3	0	0	0	0	...		3
0	.	0	0	0	4	0	0	0	...	$\mapsto$	4
0	.	0	0	0	0	5	0	0	...		5
0	.	0	0	0	0	0	6	0	...		6
0	.	0	0	0	0	0	0	7	...		7
0	.	0	0	0	0	0	0	0	...		...

Diagonal proof: The diagonal elements contain all the whole numbers, a set which is infinite but countable (given an eternity)

However, this list of decimals contains only an infinitesimal fraction of the real numbers between 0 and 1

Aside: Integers, Real Numbers, & Infinity

Cantor's Diagonal Proof that Reals are Uncountably Infinite

0	.	0	1	**1**	0	0	0	0	0	0	...		1
0	.	0	1	0	**2**	0	0	0	0	0	...		2
0	.	0	1	0	0	**3**	0	0	0	0	...		3
0	.	0	1	0	0	0	**4**	0	0	0	...	$\mapsto$	4
0	.	0	1	0	0	0	0	**5**	0	0	...		5
0	.	0	1	0	0	0	0	0	**6**	0	...		6
0	.	0	1	0	0	0	0	0	0	**7**	...		7
0	.	0	1	0	0	0	0	0	0	0	...		...

To see this, note that we can add anything we want between the decimal and this table...

These are different numbers. They also completely fill the countably infinite integers

And they are an infinitesimal slice of the real numbers between 0.01 and 0.02

Aside: Integers, Real Numbers, & Infinity

Cantor's Diagonal Proof that Reals are Uncountably Infinite

0	.	0	2	1	0	0	0	0	0	0	...		1
0	.	0	2	0	2	0	0	0	0	0	...		2
0	.	0	2	0	0	3	0	0	0	0	...		3
0	.	0	2	0	0	0	4	0	0	0	...	$\mapsto$	4
0	.	0	2	0	0	0	0	5	0	0	...		5
0	.	0	2	0	0	0	0	0	6	0	...		6
0	.	0	2	0	0	0	0	0	0	7	...		7
0	.	0	2	0	0	0	0	0	0	0	...		...

We can repeat the argument a countably infinite number of times

Producing a countably infinite number of countably infinite numbers, and never coming close to capturing all the reals

Aside: Integers, Real Numbers, & Infinity

Cantor's Diagonal Proof that Reals are Uncountably Infinite

0	.	0	3	1	0	0	0	0	0	0	...		1
0	.	0	3	0	2	0	0	0	0	0	...		2
0	.	0	3	0	0	3	0	0	0	0	...		3
0	.	0	3	0	0	0	4	0	0	0	...	$\mapsto$	4
0	.	0	3	0	0	0	0	5	0	0	...		5
0	.	0	3	0	0	0	0	0	6	0	...		6
0	.	0	3	0	0	0	0	0	0	7	...		7
0	.	0	3	0	0	0	0	0	0	0	...		...

Result: Real numbers are infinite at a level greater than integers

The real numbers cannot be counted, even given eternity

Indeed, the reals between any two real numbers are uncountably infinite

What do this have to do with POLS 205?

Finite & countably infinite variables (anything we can re-express as integers) $\rightarrow$ discrete variables

Uncountably infinite variables $\rightarrow$ continuous variables

In next five weeks, we will investigate *different* methods for each class of variable

Discrete variables

There are three types of discrete variables: Binary, Ordered, & Nominal

Binary data take on only two possible values. Without loss of generality, let these values be 0 and 1.

Examples:

Did you vote? $\{\text{No}, \text{Yes}\} \mapsto 0, 1$

Are you a Catholic? $\{\text{No}, \text{Yes}\} \mapsto 0, 1$

Discrete variables

Ordered data take on countably many values. I.e., can map to a subset of the counting numbers: $1, 2, 3, \textit{ldots}$

Examples:

Do you support 2010 Health Care reform?

$$\{\text{Does too little, Just right, Doesn't do enough}\} \quad \mapsto \quad \{1, 2, 3\}$$

Discrete variables

Ordered data take on countably many values. I.e., can map to a subset of the counting numbers: $1, 2, 3, \textit{ldots}$

Examples:

Do you support 2010 Health Care reform?

$$\{\text{Does too little, Just right, Doesn't do enough}\} \quad \mapsto \quad \{1, 2, 3\}$$

How democratic is a given country? (Polity IV)

$$\{-10, -9, \ldots 10\} \quad \mapsto \quad \{1, 2, \ldots, 21\}$$

Discrete variables

Ordered data take on countably many values. I.e., can map to a subset of the counting numbers: $1, 2, 3, \textit{ldots}$

Examples:

Do you support 2010 Health Care reform?

$\{\text{Does too little, Just right, Doesn't do enough}\} \quad \mapsto \quad \{1, 2, 3\}$

How democratic is a given country? (Polity IV)

$\{-10, -9, \ldots 10\} \quad \mapsto \quad \{1, 2, \ldots, 21\}$

How many people vote for a candidate in an election?

$\{0, 1, 2, 3 \ldots m\}$, where m is the number of registered voters

Discrete variables

Ordered data take on countably many values. I.e., can map to a subset of the counting numbers: $1, 2, 3, \textit{ldots}$

Examples:

Do you support 2010 Health Care reform?

$\{\text{Does too little, Just right, Doesn't do enough}\} \quad \mapsto \quad \{1, 2, 3\}$

How democratic is a given country? (Polity IV)

$\{-10, -9, \ldots 10\} \quad \mapsto \quad \{1, 2, \ldots, 21\}$

How many people vote for a candidate in an election?

$\{0, 1, 2, 3 \ldots m\}$, where m is the number of registered voters

How many times does the press mention "financial reform" today?

$\{0, 1, 2, 3 \ldots \infty\}$

Discrete variables

Nominal data take on name values lacking a unique ordering

Examples:

Which candidate do you prefer? $\{\text{Obama}, \text{Clinton}, \text{McCain}\}$

Which region do you live in? $\{\text{Northeast}, \text{Midwest}, \text{South}, \text{West}\}$

Discrete variables

We can't map the coding of Nominal variables to any ordering

But notice we can recode *any* discrete variable as a series of binary variables:

Which candidate do you prefer? $\mapsto$

1. Do you prefer Clinton to Obama & McCain? $\{\text{No}, \text{Yes}\} \mapsto 0, 1$
2. Do you prefer Obama Clinton & McCain? $\{\text{No}, \text{Yes}\} \mapsto 0, 1$
3. Do you prefer McCain to Clinton & Obama? $\{\text{No}, \text{Yes}\} \mapsto 0, 1$

Also notice that any two of these questions is sufficient to reconstruct the full Nominal variable

Continuous variables

A continuous variable is one that can take on any *real* value

Examples:

- Unemployment rate: Can take on any real value between 0 and 1.

Continuous variables

A continuous variable is one that can take on any *real* value

Examples:

- Unemployment rate: Can take on any real value between 0 and 1. (Or can it? Close enough?)

Continuous variables

A continuous variable is one that can take on any *real* value

Examples:

- Unemployment rate: Can take on any real value between 0 and 1. (Or can it? Close enough?)
- Gross domestic product: Can take on any positive real value. (Or can it? Close enough?)

Continuous variables

A continuous variable is one that can take on any *real* value

Examples:

- Unemployment rate: Can take on any real value between 0 and 1. (Or can it? Close enough?)
- Gross domestic product: Can take on any positive real value. (Or can it? Close enough?)
- Growth in gross domestic product: Any positive real value (Close enough?)

Continuous variables

A continuous variable is one that can take on any *real* value

Examples:

- Unemployment rate: Can take on any real value between 0 and 1. (Or can it? Close enough?)
- Gross domestic product: Can take on any positive real value. (Or can it? Close enough?)
- Growth in gross domestic product: Any positive real value (Close enough?)
- Inequality: Ratio of 90th to 10th percentile of income (Close enough?)

Notice lots of economic variables

Most political variables are discrete!

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score
Feeling Thermometer Scores

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score
Feeling Thermometer Scores

Ratio Meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, & ratio convey "factor changes"

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score
Feeling Thermometer Scores

Ratio Meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, & ratio convey "factor changes"

Examples:
Degrees Above Absolute Zero

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score
Feeling Thermometer Scores

Ratio Meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, & ratio convey "factor changes"

Examples:
Degrees Above Absolute Zero
Number of Votes Received

Additive versus Ratio scales

Continuous variables come in two flavors, depending on whether zero really means an absence of the variable:

Additive No meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, but a ratio does not.

Examples:
Degrees Fahrenheit
Polity IV Democracy Score
Feeling Thermometer Scores

Ratio Meaningful zero $\rightarrow$ "1 unit increase" has a consistent meaning across the scale, & ratio convey "factor changes"

Examples:
Degrees Above Absolute Zero
Number of Votes Received
Unemployment Rate

Exam next Wednesday, in Johnson Hall 111

Review session is Monday

Bring questions!