Descriptive findings might shape your views
A single person on welfare considered employable received $610 per month from the province. A single parent with one child received $1368 (2007).
The province spent $1.5 billion on welfare in 2012 - 3.4% of the provincial budget.
Causal findings:
There is substantial evidence that the availability of welfare benefits leads recipients to work less.

Moving across types of questions/claims
Answering one kind of questions often requires us to answer other types of questions first:
Answering prescriptive/normative questions often requires first answering descriptive and causal questions.
But: prescriptive/normative questions can never be answered with only descriptive or causal findings.
Answering a descriptive question may raise causal or prescriptive questions:
Why is the world this way?
Why should we change it?

Empirical Social Science:
Focuses mainly on causal claims
Many lead to answering descriptive questions
Prescription requires causal knowledge
Though not all causes manipulable
Prediction requires causal knowledge
Basic understanding of the work: how things are causally connected
Ever-present causal puzzles:
Why do political actors often fail to make prescriptively ideal choices?
Why do political choices and outcome vary so widely?

What is a cause?
NATO intervention caused Qadaffi's fall from power.
Means: if there had been no NATO intervention, the Qadaffi wout not have falle from power.
Weak European institutions caused the euro crisis.
If European institutions had not be weak, then the euro crisis would not have happened.
The Occupy Wall Street movement emerged because of growing inequality.
If inequality had not been growing, then the Occupy Wall Street would not have emerged.

Causal effect - Counterfactual definition
To say that C is a cause of E means: if C had not happened, then E would not have happened.

The "Arab Spring"

In December 2010 Mohamed Bouazizi, a Tunisian vegetable peddler sets himself on fire after his cars is confiscated by a policewoman.
Bouazizi self immolation ⇒ Collapse of Tunisian regime

Oppressive rule ⇒ Bouazizi self immolation ⇒ Collapse of Tunisian regime

Oppressive rule: fuel
Bouazizi self-immolation: a spark
Oppressive rule: structural cause
Bouazizi self-immolation: triggering cause.

Structural causes vs. triggering events
Triggers are details of how or when exactly an outcome happened
But triggers mostly fail to meet the counterfactual definition of a cause.
Triggers are substitutable.
Any spark will star a blaze if there is fuel.
Some trigger may be necessary, but the specific trigger is usually unimportant.
Social scientist are more interested in structural causes than in triggering events.

The cause(s) of WWI
Structural causes:
Ethno-national tensions within Austria-Hungary
Imperialism: competition for resources.
German militarism and aggressive aims.
Structure of alliances: Triple Alliance vs. Triple Entente

Structural causes vs. Individual qualities or choices
Social interaction is a sum of individual choices, but:
Individual choices often produced by structural factors
Constraints, incentives, resources, grievances
Individuals often selected for by structural factors
Individual choices often overwhelmed by structural forces.

Necessary conditions
Sometimes our causal claims take the form:
“Condition C must happen for outcome E to emerge”.
Condition C = necessary condition for E
A cause without which an effect cannot occur.
Democracy is a necessary condition for capitalism.
A middle class is a necessary condition for democracy.
War cannot happen without misperception.
There must be high levels of immigration before a government will adopt multiculturalism policies.

Conjunctural causation
When an effect depends on a combination of causes.

Multiple necessary conditions: the effect only happens when more than one condition is present.
Plant growth requite both water and sunlight.
Revolution requires both material deprivation and arbitrary rule.
War required both misperception and competition for scarce resources.
Conditional effect: the effect of cause C on E depends on cause D
2 oz alcohol causes…
A little drunkeness when person is not on medication
A lot of drunkeness when person is on antibiotics.
Misperception causes…
Short war when resources are plentiful.
Long war when resources are scarce.

Sufficient conditions
A cause that always produces and effect.
A fire is sufficient to cause heat.
Social science examples are very rare
Causation are mostly conjunctural

Multiple causation
When there is more than one set of causes that can produce an effect.
A can cause E or C can cause E
What can cause political leaders to lose power?
Military defeat
Economic decline

Multiple & Conjunctural causation
A + B can cause E or C + D can cause E
Why do civil wars happen?
Ethnic diversity + political inequality between the groups
or
Meddling by foreign powers + deep ideological polarization

Deterministic & Probabilistic causal claims
Deterministic claim: what will happen with certainty under particular causal conditions.
The effect cannot happen when the cause is absent, the effect must happen when the cause is present.
“Rapid economic growth is sufficient to cause a governing party to win the next election”
“For two countries with competing interests to go to war, at least one of them must be non-democratic”
Probabilistic causal claim
Makes an effect more likely
“Rapid economic growth makes it more likely that the governing party will win the next election”
“When two countries have competing interests, war is more likely if one of them is non-democratic”

Why social life is often more predictable?
Social interaction and choice are often patterned by:
Rules (institutions)
Goal oriented behaviour
Responsiveness to material incentives, costs, benefits
Strategic interaction
Resource constraints
Known cognitive biases

Why make probabilistic causal claims?
Imperfect predictability
Some fundamental unpredictability to human choice
Human errors
Free will
Overwhelming causal complexity
Causal factors we don't know about
May as well be random
Measurement error
All social interaction has both systematic and random components.

Specific vs General knowledge
Specific:
What caused Rwandan civil war?
Hutu-Tutsi hatred
Belgian race policies
Explains a case
General?
What causes civil war?
Ethnic diversity
Colonial race policies
Explains a phenomenon
A theory of civil war

Specific:
Why did Jane vote for NPD in 2011?
Jane is not religious
Jane is a visible minority
Jane is a recent immigrant
Explains the case of Jane´s vote in 2011
General:
What explains the size of the leftwing vote in 2011 federal election?
Religiosity
Ethnicity
Immigrant status
Explains a phenomenon in specific setting.

Even more general
What explains the size of the leftwing bote in advanced democracies?
Religiosity of population
Ethnic makeup of electorate
Immigration patterns
Explains a general phenomenon

From specific to general
What general category might we consider these case to be cases of?
Hong Kong democracy protests
Adoption of BC carbon tax
Conflict between BC government and teachers' union

The importance of general knowledge: Causal theory
Causal theory identifies systematic causes
Causal relationships that operate across space and time.
Helps make sense of the world: identifies patterns amidst complexity
Helps prescribe and predict
Also helps explain specific events
Why more democracy causes better health?
More democracy ⇒ Voting can more easily bring about removal of government ⇒ Government officials have to please larger numbers of citizens to stay in office ⇒ Governments must spend greater resources on improving citizen health ⇒ Better health of citizens.

Causal Logic
A set of statements about how and why a cause produces its effect.
Usually involves a causal chain that logically connects the cause to the effect
C ⇒ e1 ⇒ e2 ⇒ ⇒ e3 ⇒ E

Also includes assumptions:
proportional representations > more parties with different ideologies >

Why causal logics matter
More complete understanding: tells us how
Suggests other causes of the effect:
Other things that effect e1, e2, e3 will affect E
Helps with prescription
Sometimes we can't manipulate C
But we can manipulate e1, e2, e3
Make sure we manipulate C in order to get to e1, e2 or e3

Concepts and Measurement

Formulating questions and claims ⇒ Answering questions and testing claims

From fuzzy language to clear concept
A politician's success depends on the economy doing well
Politician's success
Survival in office
Economy doing well
Improvement in average standard of living

From concept to measure
Politician's length of survival in office
Improvement in average standard of living

From language to measurement
Claim with a fuzzy language
Claim with clear concepts
Measurements for those concept

Concepts and measures
- A concept:
- An abstraction used to describe characteristics of a phenomenon based on a set of criteria or qualities
- A measure
- A procedure for determining whether or to what degree a concept applies to specific cases based of observation.
- Height:
- lining a person up against a metric ruler on the wall and seeing what number is even with the top of head
- Public approval of Prime Minister Harper:
  - Ask 1000 randomly sampled Canadians what they think about the PM

‘

From concept to measure: "yardsticks"
- Improvement in average standard of living
- Measure as national income
- But we need a yardstick
  - Growth in national income
  - Growth in income per capital
  - Growth in per capita income compared to prices
  - Choosing the right yardstick for your measure is crucial
  - Depends on why you want to measure something

From concept to measure: absolute vs. relative
- Absolute measures:
  - Measures that use measurement units.
  - Do not include comparison of cases with one another

- Relative measures
  - Measures that consist of comparison of cases with one another
  - No units: percentiles, rankings, ratios

From concept to measure: categorical vs. continuous
- Categorical concepts and measures: place cases into discrete groups based on whether characteristics are present or absent
- Continuous concepts and measures: place cases along a spectrum from more or less based on a degree of measure

What are dimension of a concept?
- Part of a concept
- Could exist, disappear, change independently of one another.
  - They're not redundant with one another.
- Different dimensions, potentially **different effects**

Challenges of measurement

Measuring political knowledge
- Big debate:
  - How competent are people to engage as citizens in democratic politics?
- What do you think of these questions as measures?
  - Rainbow flag
  - Wikileaks
  - Tar sands
  - Compassionate care benefits
  - Finance minister
  - 2nd place in last election
- Two measures:
  - Ask respondents a set of factual questions; more answers right = more political knowledge
  - Have interviewer provide rating of respondent knowledgeability after 2-hour interview

Threats to measurement validity
- What can go wrong with the validity of a measure-fit with the concept?
  1. The measure does not cover **enough** of the concept.
  2. The measure covers things **outside** the concept
  3. The measure captures **different** things in different units - **non comparability**

Two criteria of good measurement
- Validity of measure
  - The degree of fit between a measure and the concept it is intended to measure

  - How well a measure “captures” the concept
- Reliability of a measure
  - How consistently a measurement procedure produces the same result when the procedure is repeated
  - If two researchers use the same procedure, do they get the same result?
  - If we use the same procedure at two different times, do we get the same result?

Political Knowledge
- Ask respondents a set of factual questions; more answers right = more political knowledge
  - More reliable, less valid
- Have interviewer provide rating of respondent knowledgeability after 2 hours interview
  - More valid, less reliable

Threats to measurement reliability
- Room for research interpretation
  - Imprecise procedures for arriving at scores
- Instability over time in the measure
  - People’s responses to survey questions change even when their underlying attitudes don’t

Hard to measure
- Concept: Racism
  - Possible Measure: ask people the following question:
    - “Do you have negative feelings toward people of other ethnicities?”
  - Problems:
    - Most people will answer “no” to not be tagged as racist.
    - The score will be lower than the right answer
- Concept: Exposure to political advertising
  - Possible measure:ask people the following question:
    - “How many campaign advertisements have you seen in the last week?”
  - Problems:
    - Keeping count
    - Relying on memory

Measurement error
- Poor validity or reliability ⇒ measurement error.

Two kinds of measurement error:
- Bias (systematic error)
  - Error produced when our measurement procedure produces scores that are, on average, either too high or too low.
  - Upward bias vs. downward bias
  - May also be too high or too low for everyone or just for certain subgroups.
  - On average: even if you repeat, you’ll tend to over or under estimate the presence of the phenomenon.

  - Not “political” bias: the researcher is not bias, it is a structural flow.
- Random error:
  - Error that derives from random features of the measurement process or the phenomenon.
  - On average, random errors cancel out over lots of iterations.

Too much to measure

- For many measurement tasks, we cannot measure all instances of a phenomenon
  - We can only take measures of a subset
- Population
  - The full set of cases (countries, individuals, wars, etc.) that we are interested in learning about.
    - Civil wars: all civil wars. Ethnic civil wars: all ethnic civil wars. Why some countries become democratic and other don’t:
- Sample:
  - The subset of the population that we actually take measures of.
    - Unemployment in Canada
      - Population: all members of the labour force
      - Sample: 56000 members of the labour force
    - And we draw an interference from the sample measure about the population

Methods of sampling

How do we know that our sample is like our population?

How do we know if our sample is representative of our population?

- Random sampling
  - Selecting cases from the population in a manner that gives every case an equal probability of being chosen
  - Law of large numbers:
    - As a random process is repeated an increasing number of times, the values generated will converge on the expected (true) value of the underlying size.
    - As the sample gets larger, we get closer to the right values.
    - As the random sample increases, the average measure taken for that sample will approach the population average.
  - Random sampling + High representativeness = more creating inferences from sample to population

Sources of error in sampling

- Sampling error:
  - The difference between what we see in the sample and the true value for the population
- Sampling bias:
  - Sampling error caused by a feature of the sampling procedure that makes some members of the population more likely to be sampled than others.
- Sampling bias sources:
  - Sampling frame ≠ population:

    - Sampling frame: group from which you draw random sample (hat from which you pick up names: phonebook)
    - Sampling bias arises if sampling frame differs from population in ways related to what you are studying. Who is excluded from the sampling frame.
    - Eg: asking people in short sleeves about favourite temperature, election poll based on random sampling from phonebook, study public agenda based on random sample of the New York Times articles. Some parts of the population is not present.
  - Self-selection:
    - Respondents often have control over whether they join your sample
    - Eg: who decides to take a survey on environmental issues?
    - Eg: internet polls based on people clicking “your opinion counts” banners.
    - Eg: who decides to take a poll about election?
      - Only 9%, which 9%?
- Random sampling error:
  - Sampling error caused by random variation between samples.
  - By pure chance, one random sample of a population will be somewhat different from another random sample of the same population
    - And different from the population average
  - Solution:
    - Increase the size of sample

Variables and Hypotheses

Variables

- A measurable property of a phenomenon that can potentially take on different values
  - Those values may be numerical or categorical: quantitative or qualitative
  - The variation may be across units or over time
    - Ethnicity (categorical)
    - Income (numerical)
    - Whether a country is at war (categorical)
    - Level of public support for prime minister (numerical)
- A way of framing a concept that emphasizes the possibility of variation across units.

Causes and effects as variables

- Causal claim: democracy increase economic growth
  - Variable: level of democracy
  - Variable: rate of economic growth
  - Independent variable: level of democracy
  - Dependent variable: rate of economic growth
  - Causal claim: a higher level of democracy lead to a higher rate of economic growth.

Independent and dependent variable

- Independent variable: the variable capturing the suspected cause in a causal claim
- Dependent variable: the variable capturing the suspected effect in a causal claim

Causal theory

We build causal knowledge by developing and testing causal theories

- A set of general claims about the cause or the effect of a class of phenomenon. A causal theory must include:
  - An independent variable
  - A dependent variable
  - A statement about what effect IV has on DV
  - Causal logic through which IV affects DV

Some notation

- X = independent variable
- Y = dependent variable

The fundamental problem of causal interference

- We never get to observe a case under the counterfactual condition; we can only observe each case under one condition
  - We observe Canada in the last decade with the liberal party suffering scandal
  - Would the conservatives have won if the liberals had no had a scandal?
    - We'll never observe this
  - We only observe every case with the level of X and Y that it actually had.

From causal theory to empirical test

- An empirical prediction:
  - A statement about what we should expect to observe if a causal claim is true
  - What we should observe if X caused Y?

A partial solution: the comparative method

- Counterfactual definition: If X had taken on a different value, then Y would have taken a different value
- We can find another case that is a lot like the case, but in which X took another value.

The comparative method: method of difference: welfare states

- Start with Canada:
  - Concentration of authority (federal levels): high (X)
  - National health insurance: present (Y)
- We don't get to rerun Canadian history with a different level of X but

A case

- A single instance of a phenomenon under investigation. A case has an outcome: it takes on one value on a dependent variable
- Eg: An election (if question is: why do some parties win and some lose?

The comparative method: method of difference

- Claim: A causes E
- Testing the claim with the comparative method.
- Suppose B and C are other potential causes of E

- Case 1: A B and C are present. E is presents
- Case 2: A is absent, B C present. Is E absent?
  - Observing this would support the causal claim.
- Crucial than other potential causes are present but not correlated with the outcome across cases
  - This rules out other explanation (B and C) supporting explanation A

- The comparative method of difference tests an empirical prediction
- Prediction to test: if we observe two cases to be the same in all relevant respects except for the value of X, then we should observe that the two cases differ in the value of Y.

Testing theories by looking for correlations
- Correlation is a clue to causation
  - One empirical prediction of most causal theories if that X and Y should be correlated if theory is true
- Correlation:
  - A relationship across cases between the values that two variables take on
  - A positive correlation: cases with higher values of X also have a higher values of Y
- Can look for correlations across large number of cases

From theory to evidence
1. Identify a phenomenon you want to understand
   - Why does this outcome happen?
   - Outcome = Effect = dependent variable (Y)
2. Develop a causal theory
   - What do you think is the cause? What is X?
   - What is the causal logic?
3. Ask yourself: how would I know if this theory was right?
   - Make an empirical prediction
   - Could be a predicted correlation between X and Y
4. Now, go and look for predicted correlation
   - Take measurements of X and Y in a set of cases
   - Is X correlated with Y across cases as predicted?
     - A test of the theory
   - If you find the right correlation, this is evidence in favour of the theory
   - If you don't find the right correlation, this is evidence against the theory.

Results of a correlational test
- X-Y correlation in right direction
  - theory strengthened
- Strong X-Y correlation in right direction
  - Theory strengthened even more
- No correlation between X-Y
  - Theory weakened
- Evidence of clear X-Y correlation, but in wrong direction

- Theory greatly weakened

Spurious correlation
- When two variables are correlated but what correlation is not a result of a causal relationship.
- Z could cause both X and Y. Therefore, X and Y are correlated, but not causally

Solutions to spurious correlation
- Include Z in the analysis
- Examine the correlation between x and Y for cases with the same value of Z
- Controlling for Z

Intervening variables
- A variable through which X influences Y
- Intervening variables do not yield spurious correlations
  - X still causes Y

Antecedent variables
- A variable that influences X
- Antecedent variables also do not yield spurious correlations as long as A does not affect Y through a pathway that does not include X.

Reverse Causation
- X does not cause Y
- Y causes X
- Its coming because causes and effects mutually reinforce one another
  - Low income - low education

The problem of randomness
- Even a perfectly random process will sometimes produce recognizable patters: apparent correlations
- Some of the correlations we see in the world are random, not real
- How can we tell the difference?

Statistics:
- Uses probability theory to tell us how likely it is that a correlation happened by chance
- Computes how closely correlated the variables are
  - When X is higher, how often is Y higher (or lower)
- Takes into account how many cases we have
  - Patterns in lots of cases > less likely to be random
  - Law of large numbers
- Tells us the probability that the correlation we see is real or random

Statistical significance
- An indicator of how likely it is that the correlation we observe is due purely to change
- A higher degree of statistical significance means its less likely to be due to chance.

•