

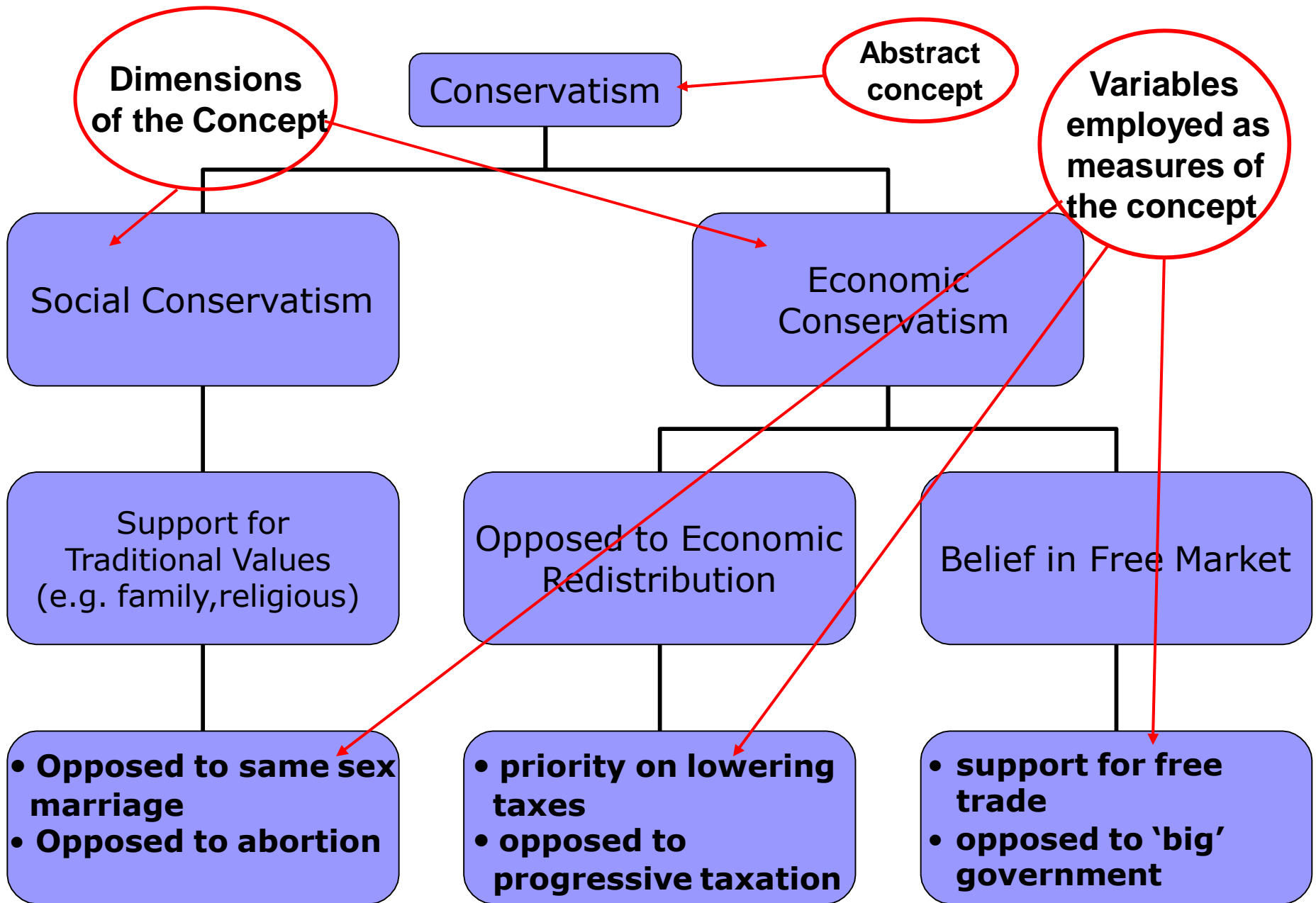
Quantitative Research: Concepts, Operationalization and Measurement

Or Getting Down to the Nitty
Gritty

Concepts to Variables to Indicators

- Research normally begins at the theoretical level
 - Concepts developed from theory, inductive method
 - Should have general agreement on what a concept means - definition
 - Conceptualization is the process of refining the definition of a concept to allow for operationalization

- Next step is to operationalize the concept
 - This step allows us to *observe* the real world
 - Move from abstract concepts to variables that can be *observed* and *measured*
 - Want achieve accuracy and precision
- Concepts are abstract; variables are concrete and specific
 - Comprehensive definition provides the framework for operationalization
- Example: operationalize ***conservatism***
 - **Def.: A political philosophy emphasizing traditional social values, classical liberal economic doctrine and opposition to radical change**



Last row are questions that would be included in a survey questionnaire

- When trying to operationalize concepts, don't use:
 - Causes
 - Consequences
 - Correlates
- For conservatism these would be:
 - Age or Income (causes)
 - Vote for the Conservative Party (consequence)
 - Religious fundamentalism (correlate?)
- Might measure these nonetheless to help ***explain*** variation in conservatism but they are not measures of conservatism itself

- Measures are simple direct items that provide data on a concept (e.g. income, age)
- Indicators, on the other hand, normally consist of several measures combined in some way to provide data on a more complex and indirect phenomenon (e.g. conservatism)
- Explaining and justifying the choices that were made in operationalizing concepts is crucial
 - Replication and transparency in research make this necessary

Operationalization – considerations when choosing measures/indicators

- Degree of variation and precision needed
- Tap all dimensions of the concept that you want to measure
 - E.g. Corruption
 - How much there is (objective – number of arrests and convictions)
 - Whether people think it exists (subjective – survey and ask people how much corruption they think exists)
 - Whether it is a problem (subjective – survey and ask people if they think it is a problem)
 - Effect on government efficiency (objective – survey politicians and public servants and ask about time spent dealing with corruption)
- Exhaustive and mutually exclusive response categories
- Levels of measurement
- Validity and reliability
- Measurement error and bias

Levels of Measurement

- Three levels of measurement:
 - Nominal (categorical) – lowest level
 - Ordinal
 - Interval – highest level
- Nominal has categories that are mutually exclusive and exhaustive but for which little else is true (e.g. place of birth, gender, religion, vote choice)

- Numbers assigned to the categories *don't mean anything in and of themselves*
 - E.g. Sex where 1=male and 5=female
- Ordinal level of measurement refers to categories for which there is an order but little else
 - E.g. Level of education where 1=did not complete high school, 2=high school graduate, 3=some post secondary, 4=post secondary degree

- Higher values indicate higher levels of the variables but the distances between categories are not always equal
- Where to break the categories depends on theory or follows established procedure
- E.g.

Please circle the response that corresponds most closely to your own opinion:

The United Nations keeps the world safe.

3 2 3 4 5 6 7 8 9

Strongly Agree Strongly Disagree

- Interval measurement includes everything that was true for ordinal and nominal variables but adds that the ***intervals between categories are equal***
 - E.g. income measured in dollar figures - \$0 means no income and \$25,000 is half as much as \$50,000; age in years; number of children in the home
 - Where a natural zero exists the interval level of measurement is referred to as ratio

- A measure can be regrouped (recoded) to form a measure at a lower but not a higher level of measurement
 - E.g. interval can become ordinal or nominal; ordinal can become nominal; Income that is measured in exact dollar amount (ratio) can be regrouped into categories of \$25000
 - 0 to \$25000 = 1
 - \$25001 to \$50000 = 2
 - \$50001 and over = 3
 - The variable is now an ordinal level variable

Validity and Reliability

- **Validity** in operationalization relates to: how well am I measuring the concept given commonly accepted definitions of it? (degree of fit, integrity)
- **Reliability** relates to: how consistent or similar are my readings across repeated measures?
- Think of hitting a bull's eye


Validity

- Want measures that are *valid and complete (measurement validity)*
 - Valid means measuring what you hope to measure
 - Complete means measuring all dimensions of a concept
- Validity also refers to research conclusions (*internal and external validity*)
- How can we assess validity?

Assessing validity


- **Face validity** exists if one can argue that there would be general agreement that the measure taps the concept/variable in question
- **Content validity** exists if the measure reflects the multiple dimension of the concept
- Face and content validity refer to how you have operationalized your measure
- **Construct validity** suggests that correctly measured concepts ought to relate in some theoretical way to other measures
 - E.g. sense of duty ought to be positively related to vote turnout

- ***Concurrent validity*** exists if different indicators of the same theoretical concept/variable yield relatively similar results across cases tested at the same time
 - E.g. most people who score high on one indicator of environmentalism should also score high on a second (that is, the scores should be positively correlated); for example, we could expect people who recycle at home to also recycle at work
- ***Predictive validity*** exists if the measure predicts an outcome at a later date according to expectations
 - E.g. children whose family discusses politics often should score higher on measures of political participation as adults

- 
- Construct, concurrent and predictive validity refer to whether the measure behaves the way you think it should (at least in a statistical sense)
 - Validity is a quality – measures will be more or less valid but not either valid or invalid
 - Using multiple indicators of a concept is always a good approach

Reliability

- Reliability deals with consistency across repeated measures
 - E.g.: it is more reliable to ask a person directly or to check their birth certificate when measuring age than it is to ask a friend or family member for this information
- Increasing the reliability of our measures *reduces the error in them*

- 
- How can we be certain that our measures are reliable?
 - Retest and check for consistency across responses
 - Verify through an external source – check reported age against birth certificate
 - Reliability increases when measure has a consistent meaning across all members of a targeted group and it is delivered objectively

Measurement Error

- When your measures deviate from the “true” value we have measurement error
- The less reliable a measure is – the more measurement error that we are likely to have
- Error can also come from less than valid indicators, respondents, interviewers, coders

- The measurement that you get will consist of “true value” (depending on validity of the measure)+ systematic error (bias) + random error (noise)
 - Goal is to try to minimize systematic and random error
- Random error differs across cases – e.g. some responses/measurements might be too high, others might be too low but there will not be a pattern in the errors

- Random error of less concern with large samples (averages out)
- Systematic error (bias) is error that is persistent and in a particular direction or way – e.g. either always high or low
 - Not a concern if you only care about the relative order of cases; is a concern if you want to say something about the overall average

Avoiding Measurement Error

- Take several measures and average out
- Use multiple indicators
- Employ random sampling
- Use sensitive measures (multiple response categories)
- Avoid confusion in wording and instructions (k.i.s.s.)
- Check data for errors

Indexes and Scales

- Indexes and scales combine the responses to several measures into a single measure
 - Better to capture many dimensions of multidimensional and complex concepts
 - E.g nationalism, alienation, feminism, conservatism
 - They provide *parsimony* – better & more information from one composite measure than from several individual ones (more variability)

- Use theory to guide variable selection
- Each variable should be discriminating
- Statistics can *help* to guide these choices
 - Correlation between items should be high (internal validity)
 - Cronbach's Alpha (measure based on strength of correlations)
 - We will return to this once we cover correlations
- Likert scales sum the scores assigned to various questions together
 - Higher scores assigned to responses that are stronger indicators of the concept in question
 - Responses range from strongly disagree to strongly agree, strongly approve to strongly disapprove

Example: Index of political participation

- Start with several measures of political activity
 - Ask people which of the following actions they have taken: written letter to a public official; signed a petition; given money to a political cause; campaigned for a party; voted; joined a protest
 - 1 point awarded for each activity they have done;
 - Index varies from 0 to 6
 - Level of political activity increases with higher scores
- Activities are not distinguished in any way

Key Terms

- Concept
- Operationalization
- Causes
- Consequences
- Correlates
- Measures
- Indicators
- Exhaustive categories
- Mutually exclusive categories
- Nominal measures
- Ordinal measures
- Interval/ratio measures
- Researcher affect
- Objectivity
- Experimenter effect
- Expectancy
- Demand characteristics
- Socially desirable or undesirable behaviour
- Data massaging
- Publication bias
- Index
- Likert Scale
- Validity
 - Measurement
 - Face
 - Construct
 - Predictive
 - Concurrent
 - Internal
 - External
- Reliability
- Measurement error
 - Systematic
 - Random