

The Elements of Research

Soc 357

Fall 2006



Part I: Concepts

1. Units of Analysis
2. Variables
3. Types of Variables
4. Operationalization
5. Reliability & Validity
6. Measurement Error

Units of Analysis

- The object(s) you want to investigate

Eg. Individuals, Households, Neighborhoods,
Organizations, Countries

Variables

- ***Characteristics*** of the objects you want to investigate; they must vary.
- Formally defined around a set of ***exhaustive*** and ***mutually exclusive*** categories
- **Always** defined by the researcher

Unit of Analysis	Variable	Categories within a variable
Individual	Income	\$ amount OR Categories: <\$10,000 \$10,000-24,999 \$25,000-39,999 Etc.
Individual	Eye color	Brown, green, blue, etc.
Household	Size (# people)	1, 2, 3, 4, ...
Organization	Gender composition (% female)	
Census tract	Average income	Mean to the nearest dollar

Types of Variables – 1

Levels of Measurement:

Nominal: Exhaustive & mutually exclusive categories

Ordinal: Categories + ranks

Interval: Ordinal + meaningful metric

Ratio: Interval + meaningful zero

Types of Variables - 2

- **Qualitative:** differences cannot be described numerically
 - Nominal & Ordinal
 - Some descriptive stats: frequency, percent, mode
- **Quantitative:** differences can be described numerically
 - Interval & Ratio
 - Mean, inferential stats

Mean, median, mode

- **Mean:** the arithmetic average
 - Add up all responses, and divide by number of respondents
- **Median:** the midpoint in a distribution
 - Half of responses are above, half below
- **Mode:** the value or category with the highest frequency

Operationalization

- You need to figure out how to measure the objects you want to study
 - Attitudes? Values? Culture? Emotional states? Power? Health?
- Formally defined as: *The rules used to assign each observation into some category of a variable*

Operationalization - 1

- The procedures to collect your data
 - Eg. Ask questions, observe
 - Question: “How often have you smoked marijuana?” VS. “Have you ever smoked marijuana?”
 - Observation: Counting number of fidgets. Film movement, count frames in which movement changes VS. face to face observation, count number of times touch head

Operationalization - 2

- Figuring out the exact distinctions/categories within your variable of interest
- If counting, how to tell the beginning and end of countable things
- If distinguishing among types of actions or characteristics, developing rules for an exhaustive and mutually exclusive set of categories.

Precision vs. Accuracy

- ***Precision***: making finer distinctions
 - Eg. Height in feet vs. inches
 - Brown eyes vs. light brown, medium brown, dark brown, hazel etc.
- ***Accuracy***: correctly classifying an observation
- A tradeoff: harder to be accurate with more precise categories

Issues with Operationalization:

- ***Reliability***: Is my measure stable and consistent?
- ***Validity***: Am I really measuring what I think I'm measuring?

Checking Reliability

- ***Inter-coder reliability***: Do two or more people agree on the categorization of an observation?
- ***Test-retest reliability***: If you use the measure over time, does it produce the same result?
- ***Split-half reliability & Internal consistency***: Do a variety of different measures produce similar results?

Checking Validity

- **Face validity:** Does it seem like a good measure?
- **Content validity:** Does the measure reflect all dimensions of what you're trying to measure?
- **Criterion-related validity:** Does your measure correspond to some other criterion for identifying differences?
- **Construct validity:** Does your measure capture the meaning of a concept, as measured in a variety of other ways?

Messy Reality - 1

- Complex environments are hard to “operationalize”
 - Eg. Zimbardo Experiment
 - Field research in general
- Use **triangulation**: multiple measures of the same phenomenon to check on validity of observations
- Use **indicators**: measures of something correlated with your variable of interest
 - Eg. “Where there’s smoke, there’s fire”

Messy Reality - 2

- How would you operationalize the following concepts in the Milgram Experiment?
 - Obedience
 - Defiance
 - Authority
 - Distress

Measurement Error

- We always assume there is EXTRA variation involved the act of measurement
- Observed value = True Value + Systematic Error + Random Error
 - Systematic error: variation caused by measurement
 - Random error: variation unrelated to measurement

Part II: Relationships

1. Independent and Dependent Variables
2. Statements about variables
3. Types of Statements
4. Statistical Associations
5. Types of Relationships

Independent and Dependent Variables

- ***Dependent***: the one the researcher is interested in explaining, aka Y
- ***Independent***: the ones that we think have some influence over the dependent variable, aka X (X1, X2, X3, etc.)
- The researcher decides which is which

Statements

- ***Relationships***: statements about two or more classes of things – people, groups, countries – that occur together and change together
- ***Proposition***: a statement about variables
- ***Hypothesis***: a statement about the expected relationship between two or more variables

Assumptions

- A proposition that taken to be true
- ***Measurement assumptions:*** your operationalization is reliable & valid
- ***Theoretical Assumptions:*** about how things generally work

Types of Statements - 1

- ***Univariate***: a statement about one variable
Eg. “Most UW students drink beer at least once a week”
- ***Bivariate***: a statement about two variables
Eg. “Male UW students drink beer more frequently than female UW students”

Types of Statements - 2

- ***Multivariate***: a complex statement about three or more variables

Eg. “Among non-depressed students, males drink beer more often than females, but among clinically depressed student males and females drink beer equally often.”

General Form of a Proposition

- ***Conceptual***: For population (P) in condition (C), independent variable (X) causes dependent variable (Y)
- ***Operational***: For sample (p) in condition (c), independent variable (x) has a statistical association with dependent variable (y)

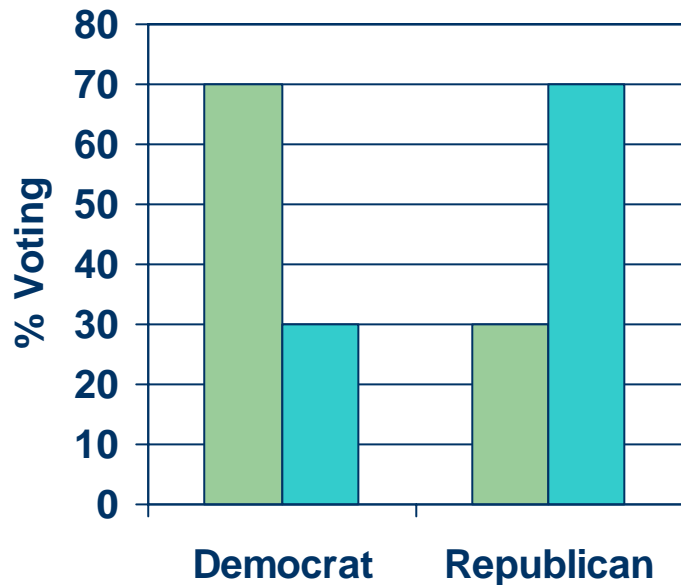
Statistical Association

- A change in one variable is associated with a change in another variable, in a way that is not likely to have occurred just by chance
- Correlations, p-values, etc.

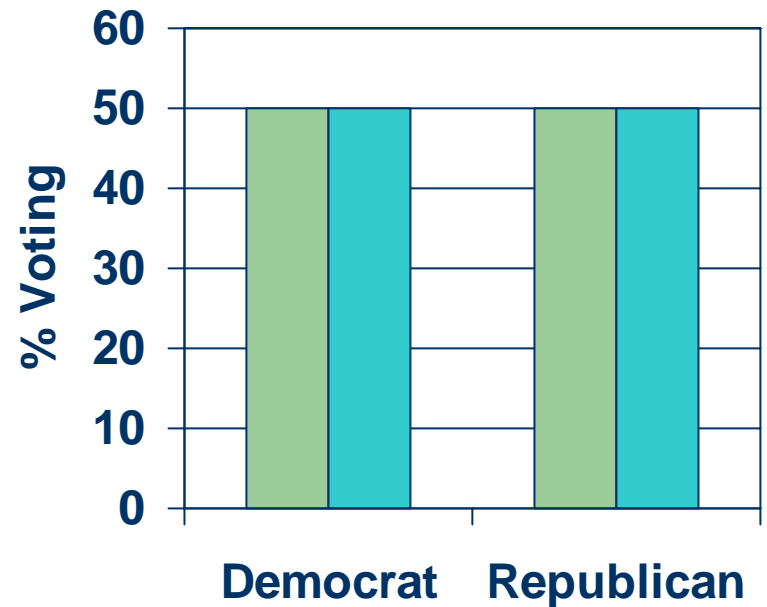
Qualitative relationships

- Among qualitative variables
- Stated in words, not numbers
- Eg. “Blacks are more likely to vote Democrat than Whites”

Qualitative relationships: Association & Non-Association



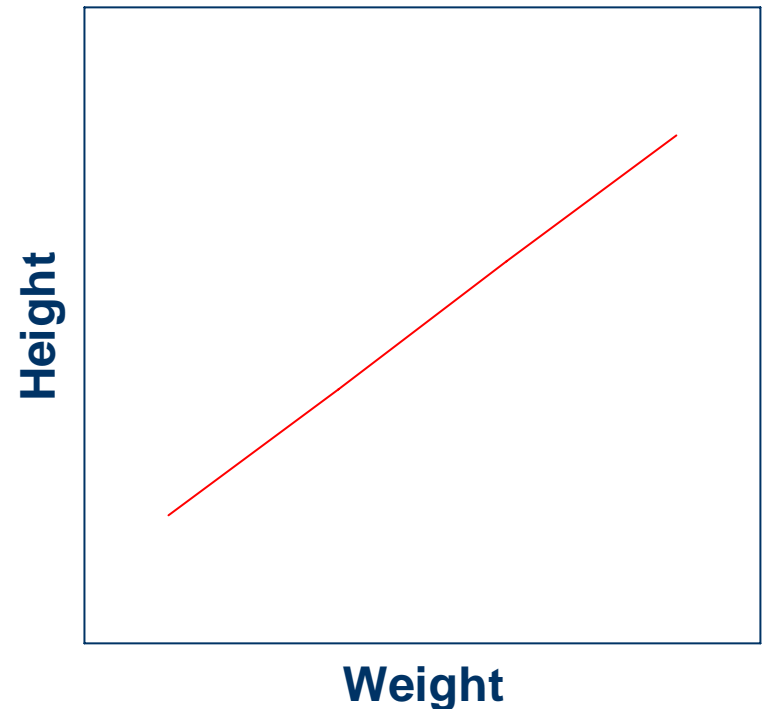
Blacks Whites



Black White

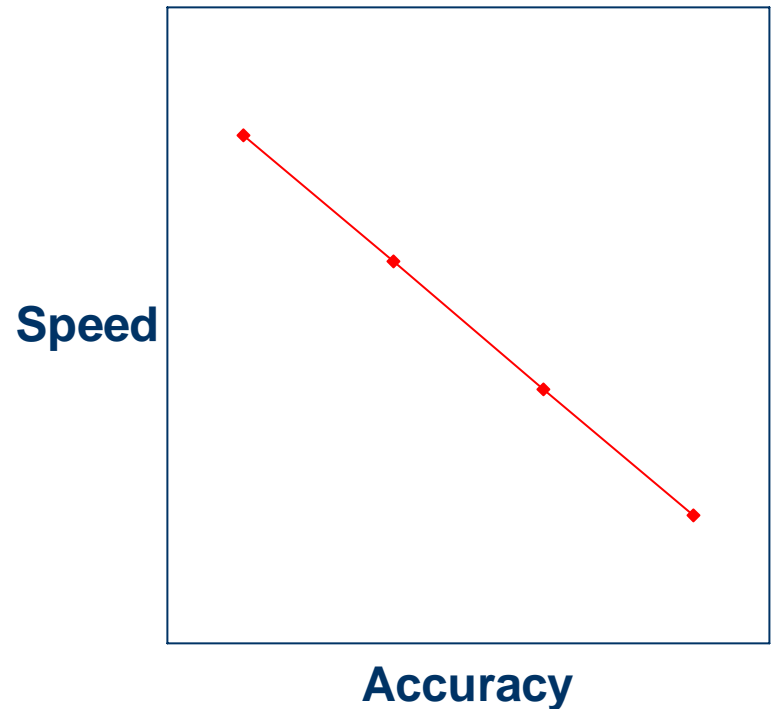
Quantitative relationships – 1

- **Positive:** when one variable is greater, the other tends to be greater
- Eg. Height is *positively* associated with weight. The taller you are, the more you are likely to weigh.



Quantitative relationships – 2

- **Negative:** When one variable is greater, the other tends to be smaller
- Eg. Speed is *negatively* associated with accuracy. The more you rush, the worse your accuracy is.



Quantitative relationships - 3

- **Curvilinear:** Any non-linear relation, but especially one that is first positive and then negative, or vice versa
- Eg. Stress is related *curvilinearly* to age. Middle aged people feel the most stress, while young & old report less stress.

