Research Design  Concepts

- Independent and dependent variables
- Data types
- Sampling
- Validity and reliability
Research Design

- Action plan for carrying out research
- ‘How’ the research will be conducted to investigate the stated problem
- Should grow directly from the reflective inquiry components
- Important research design should not dictate the reflective inquiry
Research Design

Components of the research design:

- Who or what is to be studied?
- Population or a sample?
  - Sample might or might not be representative of the population
- Variables to be measured or controlled
Research Design

- When is data collected?
- Time frame of study
- Are sampling and repeated measures used?
- Where does study take place?
Research Design

- Issues of reliability and validity
- Methodology
  - How will data be collected?
  - What is the most efficient research approach to use to collect data?
Data

Collected as a result of observation
Gathered for a purpose to support some research question
Gathered with an appropriate instrument or technique
Reliability of instrument for particular data can be measured
Susceptible to bias associated with research design
Conclusions are made based on data, systematically presented and analyzed
Research Design Concepts

- Variables and scores
  - Anecdotal data: recounting of experiences and impressions
  - Raw data: data that have not been organized or summarized
  - Variable: measurable characteristic of a population or a sample
  - Score: the value of a particular variable or an observation
Research Design Concepts

- Independent and dependent variables
- Only applicable when there is a causal or predictive relationship between variables
- Often, but not always, part of a hypothesis
- Independent determines, causes, or influences the dependent variable
- Dependent influenced or affected by independent variable
Research Design Concepts

- Independent and dependent variables examples from Powell p.31:
  - Book selection and book collection usage in academic libraries
  - The attitudes of adults toward the public library, and their relationships to library use
  - The effect of prime display location on public library circulation of selected adult titles
  - Library design influences on user behavior and satisfaction
Data Types

Any data can be considered to be one of four types:
- Nominal
- Ordinal
- Interval
- Ratio

These data types are also called scales
Research Design Concepts

Nominal data

- Classify observations into distinct classes or categories
- No inherent order among categories, though order might be used to display
- No ranking among values
- An occurrence of nominal data just falls into one category or another
- Examples are gender, marital status, political affiliation
- For analysis, values usually assigned numbers corresponding to category
  - Female = 1
  - Male = 2
- Only meaningful statistic that can be calculated is frequency of occurrence
Ordinal data

- Classify observations into ranked classes or categories
- Defines the relative position of object or individuals in a sequence, with no implication as to the distance between positions
- Also called a rank order
  - We can say that 5 is better than 4, while 4 is better than 3
  - But cannot say how much better just specifying order
- Often used to measure attitude or level of satisfaction
Research Design Concepts

- **Ordinal data continued**

  - Implemented as Likert scales, graphic rating scales, summated scales

  - Item statement 1: Strongly agree  Agree  Disagree  Strongly disagree

  - Item statement 2: Strongly agree  Agree  Disagree  Strongly disagree

  - For analysis, values assigned numbers which must be consist in order

    - Very satisfied = 5
    - Mostly satisfied = 4
    - Somewhat satisfied = 3
Research Design Concepts

- Interval data
  - Expresses order and distance between values
  - No zero point
  - Examples: common temperature scales, IQ scores
    - Celsius and Fahrenheit scales are interval because no absolute zero
    - 32°F is not twice as hot as 16°C
Research Design Concepts

- **Ratio data**
  - True zero point
  - Expresses order and distances and allows comparisons between values
  - Examples: income, age, weight
  - All ratio data is also interval data, but not all interval data is ratio data
  - Kelvin temperature scale has absolute zero $273.15 \, \text{C} = 0 \, \text{K}$
  - Both interval and ratio scale
Research Design Concepts

Sampling

Population: Set of entities in which researcher is interested, with purpose of being able to describe certain characteristics or make other predictive statements

Why not study whole population?

- Logistically impossible
- Infinite populations
- Future populations
- Destructive
- Too expensive
- Too time consuming
Research Design Concepts

Sample: Subset of a population that is examined and from which inferences are drawn about characteristics of population

Benefits of sampling:
- Reduced cost
- Results available sooner
- Broader scope yields more information
- Greater accuracy
- More attention possible to each observation

For example, compare results from expending effort to collect \( 1000 \) observations in either of the following ways:
- 1 variable observed on each of \( 1000 \) elements in the population
- 10 variables observed on each of \( 100 \) elements in a sample
Research Design Concepts

- **Sampling**
  - Goal is for each element of the population to have the same chance of being part of the sample
  - Probability sampling or random sampling increase likelihood of this happening
  - Methods available to determine population parameters and error
  - Random sample doesn’t guarantee accurate representation
  - If you toss a coin 10 times, proportion of heads will vary with each experiment (say 5)
  - If you toss coin 50 times, proportion of head likely to be closer to .5, even though same number of total tosses
  - Larger samples are better, when feasible
Research Design Concepts

Random sampling can be tricky ...
Research Design Concepts

How would you choose randomly from:

- Laboratory rats in a large cage?
- Human volunteers for a survey?
- Nuts in a large can of mixed nuts?

Is it easy to select a truly random sample in these situations?
Research Design Concepts

- Issues in selecting a truly random sample:
  - Laboratory rats in a large cage?
  - Rats will try to run away ones that are selected tend to be slowest, least active, etc.
  - Human volunteers for a survey?
  - People who volunteer likely to be different from those who don’t interested in survey topic, making money, nothing better to do
  - Nuts in a large can of mixed nuts?
  - The Brazil Nut Effect Experiment larger nuts rise to top, smaller nuts settle on bottom
Selecting a Random Sample  

Ways of selecting a simple random sample:

- Mechanical means  roulette wheel, lottery machines
  - Not especially reliable, nor practical for most research
- Toss a coin or dice, draw straws
  - Not especially reliable, nor practical for most research
- Use a printed table of random numbers  p. 72 in Powell
  - Okay for small samples, tedious for larger ones
- Use a random number generator in a computer program
  - Basically automates the printed table of random numbers approach
Research Design Concepts

- Other types of random samples
  - Systematic sample
    - Select every nth element, where each element has chance of being selected
    - For example, if 1000 elements in population and sample is to be 100, select every 10th element to be in sample
    - Elements in population should be arranged randomly in terms of characteristics relevant to study
Other types of random samples

- Stratified random sample
  - Elements in population first grouped
  - Select same percentage of elements randomly from each group proportional
  - Select same number of elements randomly from each group disproportional

- Cluster sample
  - Useful when elements of entire population cannot easily be listed for selection large populations
  - Divide population into groups clusters, select random sample from each group
  - Can occur in multiple stages census tracts, city blocks, houses, people
Research Design Concepts

Non probability sampling

- Random sample approaches are examples of probability sampling
- Each element of population has equal chance of being part of the sample
- Other types of samples do not have the same assurance
- Generally easier and cheaper to obtain
- Sometimes only feasible method of getting data
- But, prevent use of inference of findings to population
Research Design Concepts

- **Non probability sampling approaches**
  - Convenience or accidental sample
    - Selecting elements based on their availability
  - Quota sample
    - Similar to convenience sampling but select elements to match proportions found in population
  - Purposive or judgmental sample
    - Best attempt of researcher to select elements from cases already known and not considered extreme
  - Self selected sample
    - When cases volunteer or select themselves
Validity and Reliability

- Validity: are conclusions true?
- Reliability: can findings be repeated?

- Several aspects or types of each of these concepts
Validity and Reliability

Validity

- Degree to which you are truly measuring what you intend to measure
- Does the instrument measure what it is meant to measure?
- An instrument can be reliable, but not valid
  - Example: Measure anxiety with the temperature readings on a thermometer
- If an instrument is valid, it must also be reliable
Validity and Reliability

Validity

Three main types of validity:

- Internal
- External
- Construct
Validity and Reliability

**Internal validity**

- Research design should accurately identify causal relationships and rule out other explanations for relationships
- Especially important for experimental approaches
- Goal is to be sure that the conclusions drawn from experimental results accurately reflect what went on in the experiment itself (Babbie)
- Rule out possibility that confounding variables may have caused the observed results

**Threats to internal validity:**
- History, maturation, testing, compensation
Validity and Reliability

Internal validity

- Threats to internal validity continued:
  - Interaction effects: subject sensitivity to experimental variable caused by pretest, multiple treatments, experimental environment
  - Example: Hawthorne Effect study of productivity at an Electric company
  - Looking for evidence that a particular independent variable e.g. bright lights has caused a change in an observed dependent variable e.g. productivity
  - Use control groups; triangulation of methods; focus on operational definitions
Validity and Reliability

External validity

- Relates to generalizability of findings
- Even if results of study are internally valid, do conclusions hold up beyond the confines of a particular study?
- Findings should be generally true for studies conducted under a variety of circumstances or conditions other times, people, places
- Example: Study of MLIS students’ use of OPACs not generalizable to the general population, since library students have special knowledge
- Select sample carefully; repeat measure different setting/time of day/etc.
Validity and Reliability

- Construct validity
  - Properly identified and labeled variables
  - Instrument that enables accurate testing of defined variables
  - How well does the test measure comprehension of library use? Does a test on the use of an OPAC really measure effective and efficient use rather than one’s ability to read test items?
Validity and Reliability

Reliability

If the design of a research study is reliable, then its findings should be repeatable, replicable, generalizable.

- Can the study be replicated?
- Will the research yield stable, consistent results when applied repeatedly?

Example: “How many books have you borrowed this year?”

- A study in which this is an important question might be unreliable; subjects likely will not recall the exact number, will guess different numbers at different times.
Validity and Reliability

- Measurement of reliability
  - Based on a correlation between two events
  - Example: Measure reliability of test scores
    - Administer test to same group of subjects on successive days with no intervening instruction
    - If test scores are reliable, each subject will have very similar scores on each test
    - Closer the pairs of scores are /parenleft standard high correlation /parenright, the more stable or reliable they are
Validity and Reliability

- Ensuring reliability
  - Pilot studies
  - Interviewer training
  - Operational objectives and well defined variables so that measures are applied consistently
  - Same data collection instrument to observe or collect scores
Validity and Reliability

- Reliability
  - Example research question:
    - Does the library have adequate facilities and equipment for physically disabled students?
  - What are some weaknesses related to reliability?
Validity and Reliability

Reliability

- Example research question:
  - Does the library have adequate facilities and equipment for physically disabled students?
  - What are some weaknesses related to reliability?
  - Can patrons in wheelchairs retrieve books from the browsing collection?