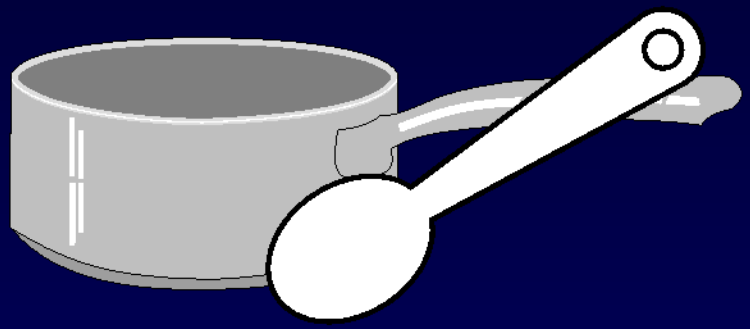
IPDET

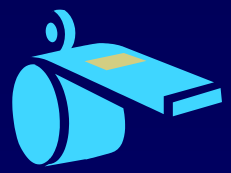
Module 7: Selecting Designs for Cause-and-Effect, Normative, and Descriptive Evaluation Questions

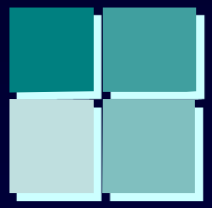
Introduction

- Connecting Questions to Design
- Design for Cause-and-Effect Questions
- Designs for Descriptive Questions
- Designs for Normative Questions
- The Need for More Rigorous Designs

What Makes Elephants Go Away?

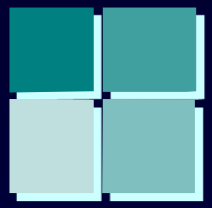






Connecting Questions to Design

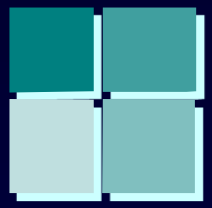
- Design is the plan to answer evaluation questions
- Each question needs an appropriate design
- Avoid the “method in search of an application” technique
- No one best design



Experimental Design

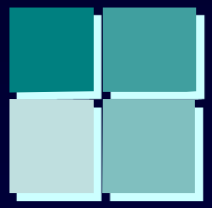
- Randomized or true experimental design

  Uses two groups, one receives intervention, other group, called the control group, does not
- Assignment to groups is random



Quasi-Experimental Design

- The design is similar to true experimental design but:
  - no random assignment
  - uses naturally-occurring comparison groups
  - requires more data to rule out alternative explanations

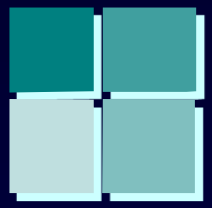


Nonexperimental Design

- Do not compare groups
- Provide extensive descriptions of the relationship between an intervention and its effects
- Evaluator attempts to find a representative sample
- Might analyze existing data or information
- Looks at identifying characteristics, frequency, and associations

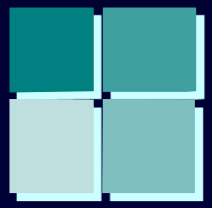
Comparison of Design Categories

Type of evaluation design	Randomly assigned control group	Non-random comparison group	Repeated measures
Experimental	Yes	No	Yes
Quasi-Experimental	No	Likely	Likely
Nonexperimental	No	No	No



Design for Cause-and-Effect Questions

- Can use experimental and quasi-experimental designs
- Pose the greatest methodological challenges
- Need a well thought out design
- Design attempts to rule out feasible explanations other than the intervention
- “What would the situation have been if the intervention had not taken place?”



Steps in Experimental Design

- Formulate a hypothesis
- Obtain a baseline (measure the dependent variable)
- Randomly assign cases to intervention and nonintervention (control) group
- Introduce the treatment or independent variable in the intervention
- Measure the dependent variable again (posttest)
- Calculate the differences between the groups and test for statistical significance

Experimental Design Notation

O_1 X O_2

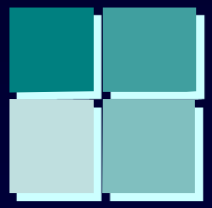
O_1 O_2

---

or

R O_1 X O_2

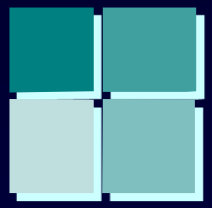
R O_1 O_2



Control Groups

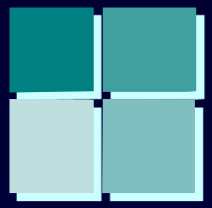
- Control group: group whose members are NOT exposed to or provided an intervention
- Treatment group: group whose members are exposed to or provided an intervention

Alternative explanations *must* be ruled out before drawing conclusions



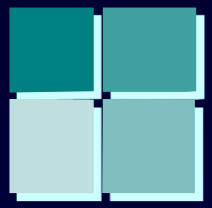
Random Assignment

- Random: people or things are placed in groups by chance
- Random assignment is assumed to make groups comparable
- Not always an option but it is possible more often than you think
  - when not all participants can receive the intervention at once



Selection Bias

- Distortion of evidence or data about the results of a program intervention due to systematic difference in the characteristics of the subset of the population receiving the intervention and those in the same population not receiving the intervention
  - self-selection of participants
  - program managers select participants most likely to succeed



Internal Validity

- Ability of a design to rule out all other potential alternate factors or explanations for the observed results other than the intervention

Common Threats to Internal Validity

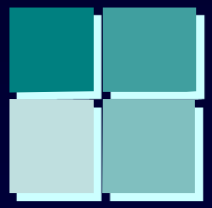
- History (events occurring at the same time)
- Maturation of subjects (getting older changes the results)
- Repeated testing (learning how to take the test)
- Selection bias (participants may be different to begin with)
- Mortality (participants departing)
- Regression to the mean (scores at extremes)
- Instrumentation (changes in data collection instruments or procedures)

Quasi-Experimental Designs

- Similar to experimental design, but does not randomly assign individuals to groups
- Compares groups that are similar but not equivalent
- When not possible to randomly assign, need to construct comparison groups
- Without random assignment, must control for internal validity

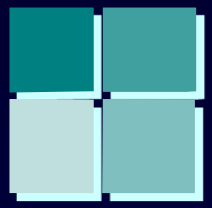
Quasi-Experimental Designs

- No real control group, evaluator constructs treatment and comparison groups by:
  - constructing groups that are equivalent on important characteristics:
    - age, gender, income, socioeconomic background, etc.
  - finding a comparison group by matching key characteristics



Examples of Quasi-Experimental Design

- Before-and-after design without comparison group
- Pre- and post-nonequivalent comparison design
- Post-only nonequivalent comparison design
- Interrupted time series comparison design
- Longitudinal design
- Panel design
- Correlational design using statistical controls
- Propensity score matching



Before-and-After Design without Comparison Group

- One way to measure change
- Compare key measures before and after the intervention
- Also called pre- and post-designs
- The before is often called the *baseline*
- There is no separate comparison group, the “before” is one group and the “after” is the same group
- Change alone does not prove causality

Questions for Before-and-After Design

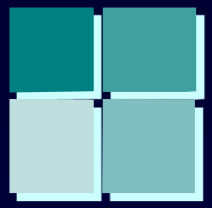
- Evaluation questions:
  - Did program participants increase their knowledge of parenting techniques
  - What was the change in wages earned, two years after the training intervention

Notation for Before-and-After Design

- Represented as:

  O_1 X O_2

  - observation, intervention, observation



Pre- and Post-Nonequivalent Comparison Design

- To make groups as similar as possible, match them using:
  - skill tests, performance tests, judgment scores, etc.
- Each subject gets scored, then place subjects in groups matching scores
  - subjects are assigned by scores, similar number of high, middle, and low scores

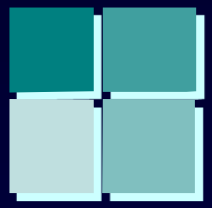
Example of Pre- and Post-Nonequivalent Comparison Design

- Evaluating gender awareness training
- Construct two groups:
  - give pre-test
  - from scores on pre-test
    - place one of the two highest scores in one group, the other in the second group
    - place one of the next highest score in one group and the other in the second group
    - etc.
  - designate one group as treatment, the other as control

Notation for Pre- and Post-Nonequivalent Comparison Design

N O_1 X O_2

N O_1 O_2



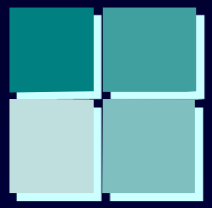
Post-Only Nonequivalent Comparison Design

- Weaker design than pre and post nonequivalent comparison design
- Comparison group exists, but there are data only for post-intervention
- Know where groups ended, but not where they began

Notation for Post-Only Nonequivalent Comparison Design

N O_1 X O_2

N O_2



Interrupted Time Series Comparison Design

- Look for change over time
- Purpose is to explore and describe changes over time – either after, or before and after the intervention
- Can be used to discern trends
- Often there are existing data that can be used

Questions for Interrupted Time Series Comparison Design

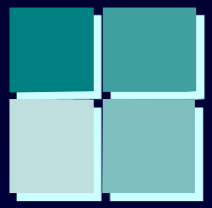
- Evaluation questions:
  - What are the trends in deminer accident rates before and after and over time for an intervention?
  - What are the changes in participant attitudes towards mine survivors as entrepreneurs after the survivors completed a three month workshop?

Notation for Interrupted Time Series Comparison Design

Represented as:

$$O_1\ O_2\ O_3\ \ X\ \ O_4\ O_5\ O_6$$

at least three observations are made prior to the intervention and again three more times after the intervention



Longitudinal Design

- A type of time series design that occurs over a long period of time
- Repeated measures of the same variable are taken from the study population
- Can give a wealth of information
- Diminishing numbers over time as subjects die or move out of contact

Example for Longitudinal Design

- Evaluation question:
  - How did the provision of micro credit affect the transition of households into and out of poverty?
    - a study looking at a landmine survivor micro credit program in Cambodia

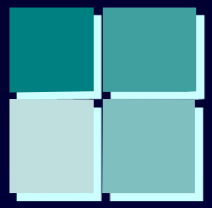
Notation for Longitudinal Design

$X \quad O_1 \quad O_2 \quad O_3 \quad \ldots Ox$

intervention followed by observations of an individual or group over time

Panel Design

- Panel design is one type of longitudinal study where a group of people are tracked at multiple points over time
  - almost always use qualitative questions (open-ended survey questions, in-depth interviews, and observation)
  - can give a more in-depth perspective



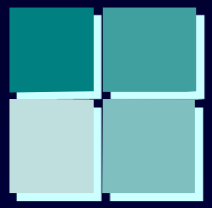
Example of Panel Design

- Investigating attitudes and patterns of behavior about gender for students at a school
- Questionnaire given every semester for eight years

Notation for Panel Design

$X \quad O_1 \quad O_2 \quad O_3 \quad O_4 \quad O_5 \quad O_6 \ldots$

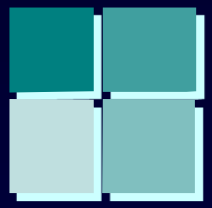
intervention followed by observations of a unit over time



Correlational Design Using Statistical Controls

- Looks at variables that cannot be manipulated
- Each subject is measured on any number of variables and statistical relationships are assessed among the variables
- Data analyst usually analyzes the data

(continued on next slide)

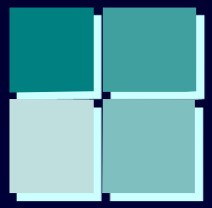


Correlational Design Using Statistical Controls (cont.)

- Often used when seeking to answer questions about relationships and associations
- Often used with already available data

Example of Correlational Design

- Evaluation looking for link between occupation and the incidence of UXO accidents
- Distribute questionnaire to large percent of the population
- Ask questions about:
  - occupation, who they contact, where they spend time away from home, etc.

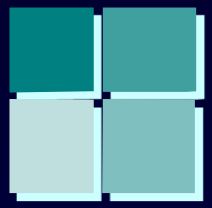


Propensity Score Matching

- Used to measure an intervention's effect on program participants relative to nonparticipants with similar characteristics
- Collect baseline data then identify observable characteristics that are likely to link to the evaluation question

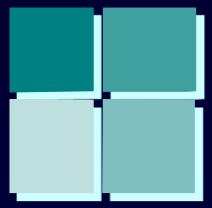
Example Propensity Score Matching

- For example, match
  - gender, age, marital status, distance from home to school, room and board arrangements, number of siblings, graduating from secondary school, birth order
- Result is pairs of individuals or households that are as similar to one another as possible (except on treatment variable)



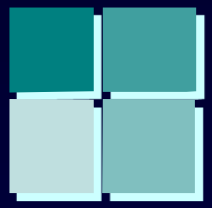
Nonexperimental Designs

- Simple cross-sectional design
- One-shot design
- Causal tracing strategies
- Case study design



Simple Cross-Sectional Design

- Show a snapshot at one point in time
- Also interested in sub-group responses
- Often used with survey method
- Subgroups may be:
  - age
  - gender
  - income
  - education
  - ethnicity
  - amount of intervention received



Questions for Simple Cross-Sectional Design

- Evaluation question may focus on
  - participant satisfaction of services
  - why they did not use services
  - find out current status of people from an intervention a few years ago
- Evaluation questions might be:
  - Do participants with different levels of education have different views on the value of training?
  - Did women receive different training services than their male counterparts?

Notation for Simple Cross-Sectional Design

Represented as:

X O_1
O_2
O_3

“

the observation is made after the intervention “X” and responses of subgroups (“O_1, O_2, O_3” and so on) receiving the interventions are examined

One-Shot Design

- A look at a group receiving an intervention at one point in time, following the intervention
- Use to answer questions such as:
  - How many women were trained?
  - How many participants received job counseling as well as vocational training?
  - How did you like the training?
  - How did you find out about the training?

Notation for One-shot Design

Represented as:

$$X \quad O_1$$

- there is one group receiving the treatment "X" and one observation "O"
- there is no before treatment / intervention measure

Causal Tracing Strategies

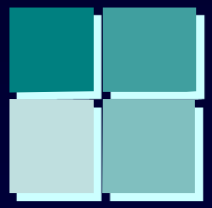
- Based on the general principles used in traditional experimental and quasi-experimental designs, but:
  - can be used for rapid assessments
  - can be used without high-level statistical expertise
  - can be used on small scale interventions where numbers preclude statistical analysis
  - can be used for evaluations with a qualitative component
  - involves the evaluator doing some detective work

Causal Tracing Strategies

- Ask yourself:
  - What decisions are likely to be based on the evidence from this evaluation?
  - How certain do I need to be about my conclusions?
  - What information can I feasibly collect?
  - What combination of information will give me the certainty I need?
- Remember: this list is a menu of possible sources of evidence, not a strict checklist of requirements

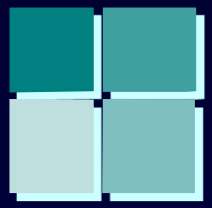
9 Causal Tracing Evidence Sources

- Causal list inference
- Modus operandi
- Temporal precedence
- Constant conjunction
- Contiguity of influence
- Strength of association
- Biological gradient
- Coherence
- Analogy



Case Study Design

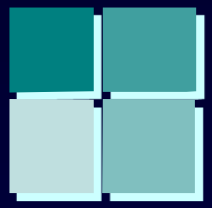
- Descriptive case study
- In-depth information is collected over time to better understand the particular case or cases
- Useful for describing what implementation of the intervention looked like – and why things happened the way they did
- May be used to examine program extremes, or a typical intervention



Notation for Case Study

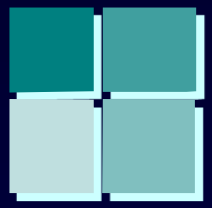
- Represented as:

  O_1
  O_2
  O_3



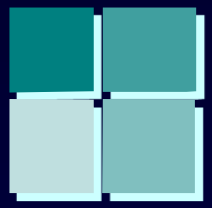
Case Study Design

- Used when the researcher wants to gain an in-depth understanding of a process, event, or situation
- Good to learn how something works or why something happens
- Are often more practical than a national study
- Can consist of a single case or multiple cases
- Can use qualitative or quantitative methods to collect data



Designs for Descriptive Questions

- Descriptive questions generally use nonexperimental designs
- Common designs for descriptive questions:
  - simple cross-sectional
  - one-shot
  - before-and-after
  - interrupted time series
  - longitudinal
  - case studies

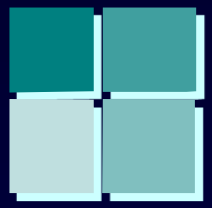


Designs for Normative Questions

- Similar to descriptive questions
- Normative always assessed against a criterion:
  - a specified desired or mandatory goal, target, or standard to be reached
- Generally the same designs work for normative questions as descriptive questions

The Need for More Rigorous Designs

- Greater call to demonstrate impact
- Evaluations should strive for more rigor in evaluation design
- Advice:
  - build the design on a program theory model
  - combine qualitative and quantitative approaches
  - make maximum use of available secondary data
  - if possible, include data collection at additional points in the project cycle



Making Design Decisions

- There is no perfect design
- Each design has strengths and weaknesses
- There are always trade-offs – time, costs, practicality
- Acknowledge trade-offs and potential weaknesses
- Provide some assessment of their likely impact on your results and conclusions

A Final Note….

“Design is not just what it looks like and feels like. Design is how it works.”

-- Steve Jobs

Questions?