RESEARCH METHODS

TYPES OF EVALUATION DATA

Qualitative

- Interviews
- Observations
- Focus Groups

Quantitative

- Surveys/Questionnaires
- Records
- Statistics

TYPES OF DATA COLLECTION

- Qualitative Research
  - Focus groups
  - Interviews
- Quantitative Research
  - Surveying
  - Experimentation

QUALITATIVE RESEARCH

FOCUS GROUPS

- In-depth discussions on a specific topic led by a moderator
- Elements:
  - 6-12 people who talk to the moderator and each other
  - Participants are paid for their time
  - Usually 1-2 hours together

FOCUS GROUPS

- 1. Define the problem
- 2. Select a sample
- 3. Determine the # of groups
- 4. Prepare the study materials
- 5. Conduct the session
- 6. Analyze the data
- 7. Prepare a summary report

delve®

ADVANTAGES OF FOCUS GROUPS

- Great for testing new ideas and getting reactions
- Can be conducted quickly
- Relatively inexpensive
- Flexible question design
- Responses are more complete and less inhibited than one-on-one interviews

DISADVANTAGES OF FOCUS GROUPS

- Cannot generalize the findings
- Participants can be swayed by a dominant participant
- Moderators must be professional and objective; if not, it could impact results

INTENSIVE INTERVIEWS

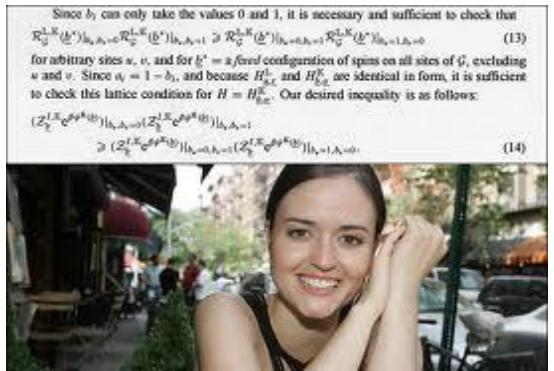
- Meant to be interactive
- Types
  - **Structured:** Specific sets of questions
  - **Open:** Develop new questions as you go along
  - **Depth probing:** "tell me more," "please explain"

DEVELOPING INTERVIEW QUESTIONS

- Hugely important step!
- Research questions ≠ interview questions
  - Interview questions are more specific and contextual
  - Should be based on literature and observations
- Do NOT ask leading questions
  - Start broadly

EXAMPLE INTERVIEW QUESTIONS

- RQ: What types of barriers do women have to overcome in male dominated majors?
  - What led you to pick your major?
  - What struggles have you faced in your major?
  - How do the male students in the department treat you?



MORE ON INTERVIEWING

- May take several sessions
  - Try to let interviewee select location and time
- Be aware of the power dynamic
- Learn to be comfortable with silence
- Listen analytically
- Interact using neutral regard

QUANTITATIVE RESEARCH

SURVEY METHODOLOGY

WHAT IS SURVEY RESEARCH?

- Systematic collection of data from individuals
  - Gather info about thoughts, feelings, and behaviors
- Can identify correlations (relationships) between variables
- CANNOT determine causation

TYPES OF SURVEYS

- Self-administered
  - Online
  - Paper
- Face-to-face
- Telephone

SELF-ADMINISTERED SURVEYS

- Anonymous
- Reflection possible
- No interviewer bias
- Not bound by geography
- Relatively inexpensive

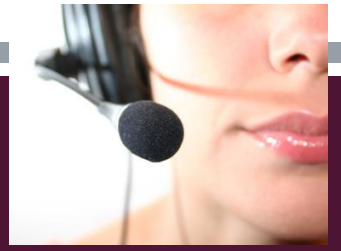
- Requires computer literacy (if online)
- Low response rate
- Must be self-explanatory
- Can't be too long
- No control over who responds

FACE-TO-FACE SURVEYS

- Can get richer info
- Responses are usually honest
- Interviewers can build rapport

- Expensive
- Recruitment is challenging
- Potentially dangerous
- Potential for interviewer bias

TELEPHONE SURVEYS



- Broad geographic reach
- Rapid data collection
- Doesn't require literacy

- Responses may not be well thought out
- High refusal rate

TIPS FOR WRITING SURVEY QUESTIONS

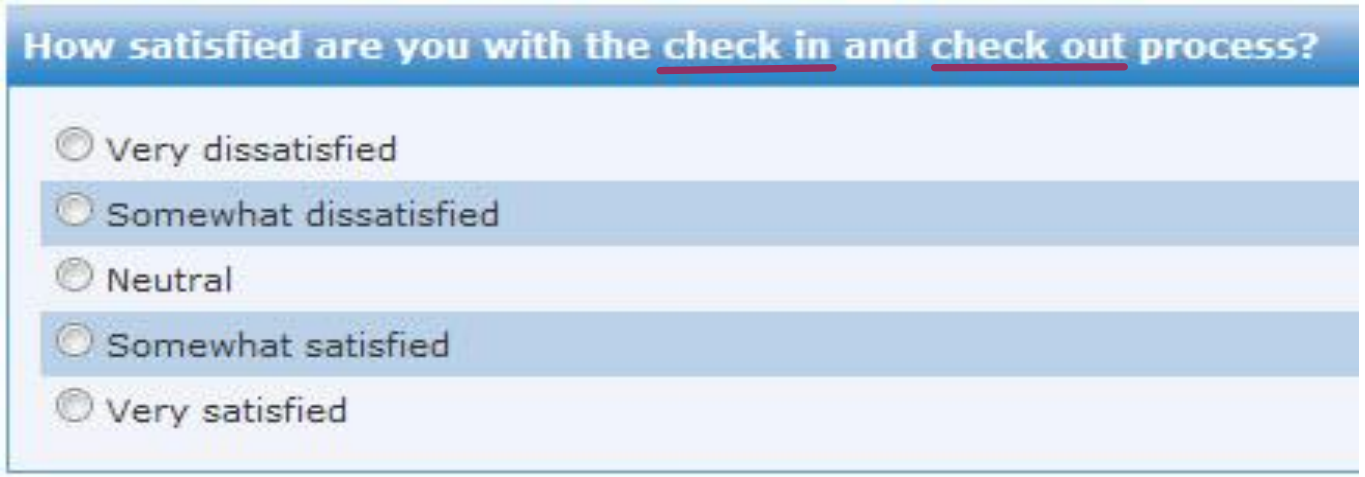
MAKE QUESTIONS CLEAR – BE PRECISE

- Avoid jargon and pretentious language
- Be fair to all types of respondents
- Keep questions short
  - Better chance of being understood

	True	False
I would definitely purchase a data plan that would allow users on my family share plan to share a fixed amount of Mobile Data (e.g., 4GB) each month for a fixed price.	○	○

AVOID DOUBLE-BARRELED QUESTIONS

- Double barreled questions: Ask about more than one thing in a single question



AVOID DOUBLE BIND QUESTIONS

- Double bind questions: Force an answer
  - Basically a "trick question" – for example:

    "It's like if someone asked you, 'You don't want dessert, right?' But you do want dessert so you say, 'Yes,' which really means you don't want dessert. And if you say, 'No,' which means you do want dessert - it sounds like you don't. Either way, you don't get what you want." – Ellen Degeneres

AVOID BIASED WORDS/TERMS

- Question phrasing can influence responses
  - “In your free time, would your rather read a book or just watch television”

AVOID BIASED WORDS/TERMS

- Question phrasing can influence responses
  - "In your free time, would your rather read a book or ***just*** watch television"

AVOID BIASED WORDS/TERMS

- Question phrasing can influence responses
  - "In your free time, would your rather read a book or ***just*** watch television"
  - "Where did you hear the news about Osama bin Laden's death?"

AVOID BIASED WORDS/TERMS

- Question phrasing can influence responses
  - "In your free time, would your rather read a book or ***just*** watch television"
  - "Where did you ***hear*** the news about Osama bin Laden's death?"

AVOID LEADING QUESTIONS

- Leading Questions: Question wording suggests a certain response

How likely are you to stop using Facebook in the next month in reaction to its recent user privacy violations?

- Very likely
- Likely
- Unlikely
- Very unlikely

DON'T ASK FOR TOO MUCH DETAIL

- Questions that ask for highly detailed information are rarely answered accurately
  - "On an average week, how many hours do you spend watching television?"
  - "In the past month, how many times have you dined out or ordered takeout?"

TRY NOT TO ASK EMBARRASSING QUESTIONS

- Very important for face-to-face and telephone surveys
- If participants feel uncomfortable, they likely won't respond thoroughly/honestly, if at all

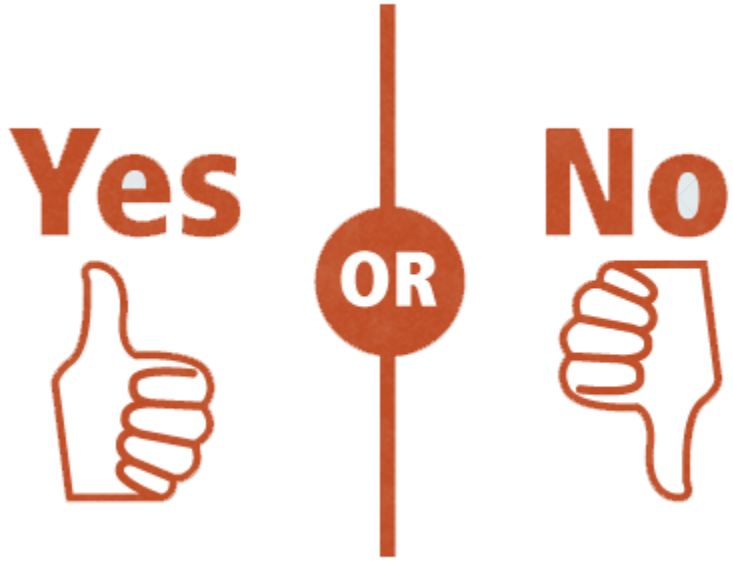
TYPES OF QUESTIONS

OPEN-ENDED QUESTIONS

- Cannot be answered with a simple "yes" or "no"
- Encourage a full, meaningful answer
  - Ex: "What would you do if…"
  - Ex: "How do you feel about…"
- Downsides? Take longer to collect and analyze

1. DICHOTOMOUS RESPONSE

- Response is one of two contradictory choices. For example:
  - "Yes" or "No"
  - "Agree" or "Disagree"
  - "Good" or "Bad"



2. MULTIPLE CHOICE

- Choose from several response options
  - Must be mutually exclusive
    - One response possible per person
    - Response options can't overlap
  - Must be inclusive
    - All possible responses listed

Select your eye color:

Blue	Green
Black	Hazel
Brown	Amber
Gray	Pink
Dichromatic	

WHAT'S WRONG WITH THIS?

- How many years have you been employed at ODH?
  - 0 to 1 years
  - 1-5 years
  - 5-10 years
  - 10+ years

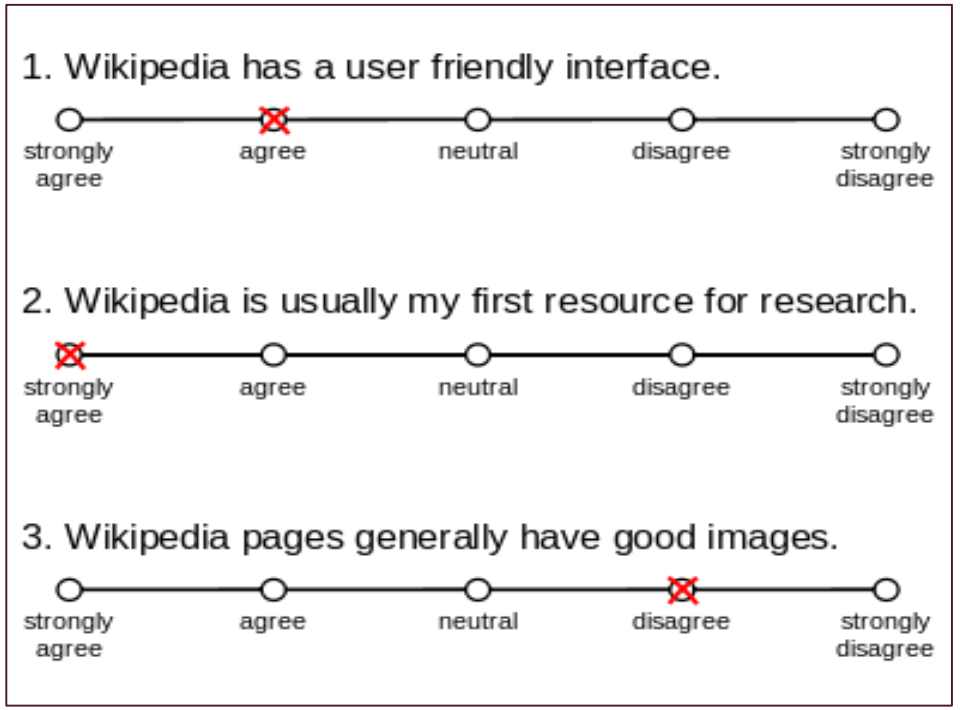
Response options overlap

WHAT'S WRONG WITH THIS?

- How many years have you been employed at ODH?
  - 0 to 1 years → Less than one year
  - 1-5 years → 1-5 years
  - 5-10 years → 6-10 years
  - 10+ years → 11+ years

3. LIKERT SCALES

- Respondents rate their agreement with an item
- Most common type of scale used



DECIDING ON RESPONSE OPTIONS

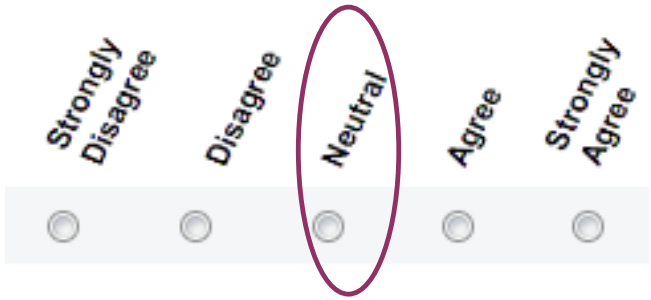
- Usually around 5 to 7 response options, but this can vary
  - Too few? Not enough nuance.

  - Too many? Too much nuance

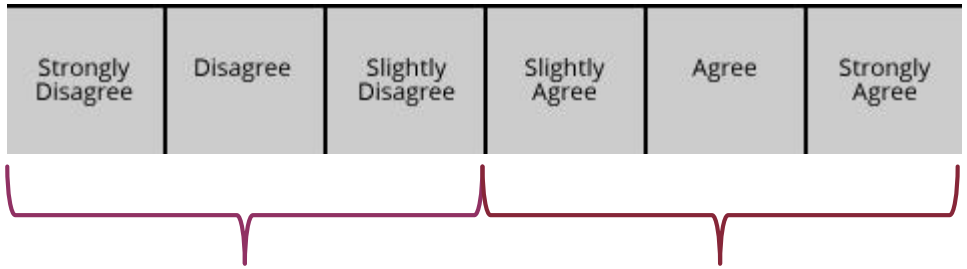
Strongly Disagree	Disagree	Moderately Disagree	Mildly Disagree	Undecided	Mildly Agree	Moderately Agree	Agree	Strongly Agree

DECIDING ON RESPONSE OPTIONS

- Odd number: Allows for a "neutral" response



- Even number: "Forced" response



DECIDING ON RESPONSE OPTIONS

- Try to use the same response options for a group of questions
  - Makes analysis and interpretation easier
- The ends of the scale are called the "anchors"
  - Don't change the polarity (i.e., keep them in the same direction)

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	2	3	4	5

SEMANTIC DIFFERENTIAL SCALES

- Two bipolar (opposite) adjectives separated by spaces:

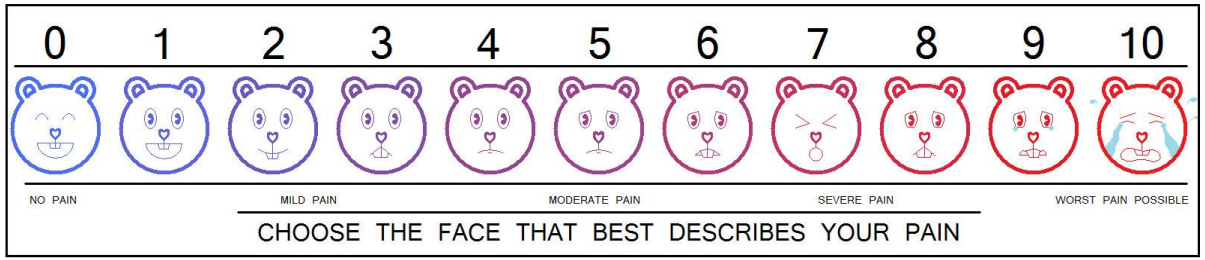
How do you feel about the recent dress code changes?

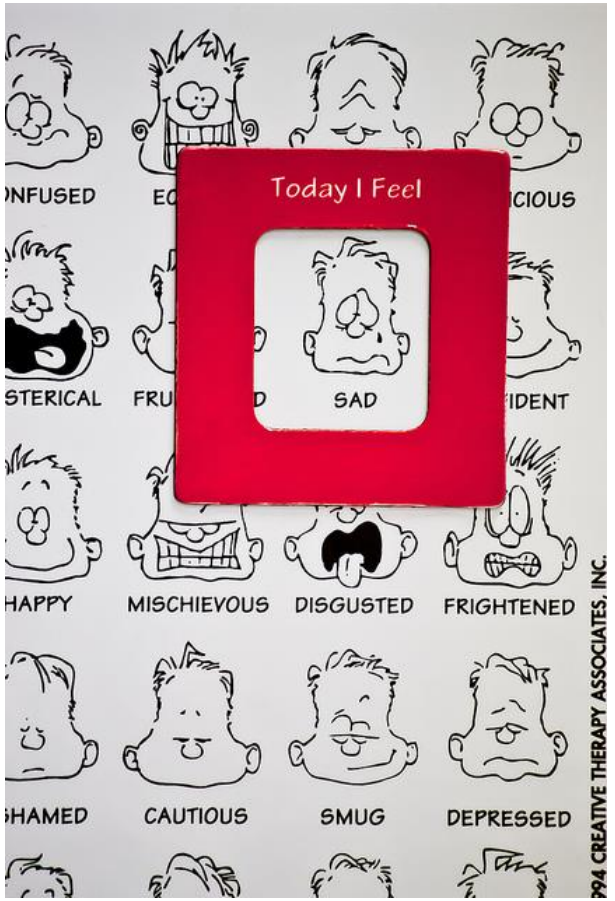
Annoyed __ __ __ __ __ __ __ Pleased

Good __ __ __ __ __ __ __ Bad

PICTORIAL RATING SCALES

- Use pictures to rate responses
- Often used with children





RANK ORDERING

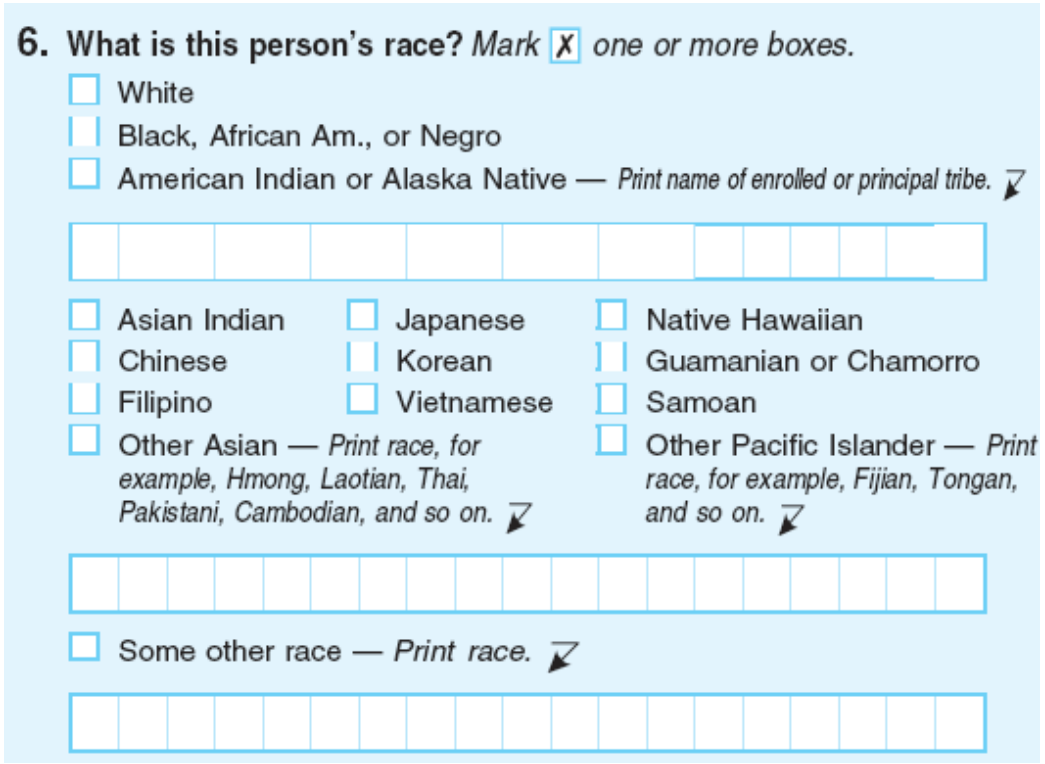
- Participants rank items based on personal perceptions and/or preference

Please rank the following from 1 to 5 according to their importance. 1 is most important.

Price	2
Comfort	1
Ease of Use	3
Stylish	5
Durability	4

4. CHECKLIST QUESTION

- Participants can select more than one option
  - Often good for participant demographics



6. **What is this person's race?** *Mark* ✗ *one or more boxes.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native — *Print name of enrolled or principal tribe.*

- ☐ Asian Indian
- ☐ Chinese
- ☐ Filipino
- ☐ Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*
- ☐ Japanese
- ☐ Korean
- ☐ Vietnamese
- ☐ Native Hawaiian
- ☐ Guamanian or Chamorro
- ☐ Samoan
- ☐ Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.*

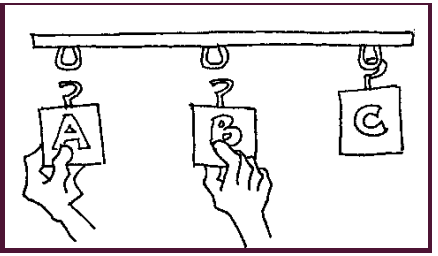
- ☐ Some other race — *Print race.*

5. FILL-IN-THE-BLANK

- Used infrequently, but good for testing recall
  - "The main topic of the webinar was _____."

OTHER CONSIDERATIONS

QUESTION ORDER



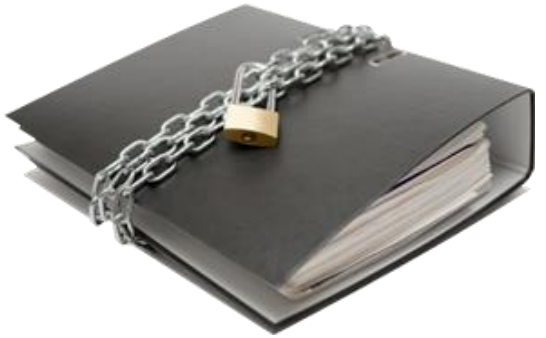
- Ask relevant, but easy questions first
  - Get participants comfortable/engaged
- Ask demographic/personal questions last
- Organize logically
  - From general to specific
  - Group similar questions together

QUESTIONNAIRE LENGTH

- Shorter questionnaires = higher response rates
- Long questionnaires cause:
  - Respondent fatigue
  - Low completion rates

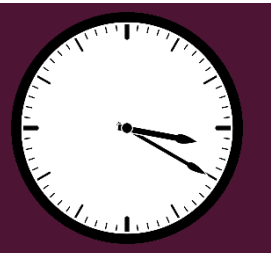
MISCELLANEOUS TIPS

- Explain the purpose of the survey
  - Provides motivation to participants
- Assure confidentiality
  - Responses are securely stored and won't be identifiable



- Have multiple people test your survey
- Consider offering an incentive (e.g, raffle for giftcard)

TIMING



- Get your surveys done ahead of time so they are ready to go
  - Especially important if it is a post-event survey
- Send within a few days of an event
- If doing a follow-up survey (e.g., 3 months later):
  - Make sure you ask about things they can actually remember
    - Don't focus on the content from the event
  - Ask about implementation
    - Changes made, number of people served, etc.

QUANTITATIVE RESEARCH

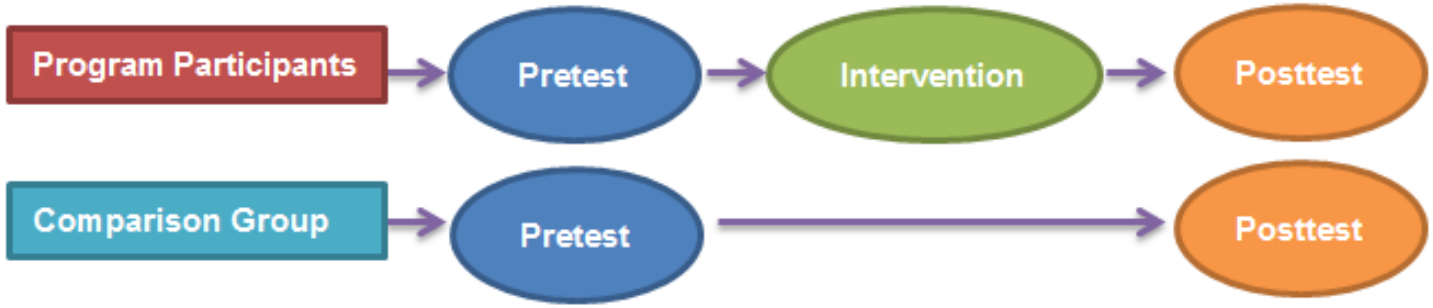
QUASI-EXPERIMENTATION

WHAT IS AN EXPERIMENT?

- A way to test cause and effect, such as:
  - "A causes B"
    - Ex: "Smoking causes cancer"
  - "As A increases, B decreases"
    - Ex: "As exposure to sunlight increases, depression decreases"

QUALITIES OF AN EXPERIMENT

- Researcher manipulates only one variable
- Two groups of participants, which are randomly assigned:
  - Experimental group – Receives the manipulation
  - Control group – Does not receive the manipulation
- Three steps:
  - Pretest
  - Intervention
  - Posttest



WHAT IS A COMPARISON OR CONTROL GROUP?

- A group of individuals not participating in the program or receiving the intervention
- Necessary to determine if the program, rather than some other factor, is causing observed changes
- "Comparison group" is associated with a quasi-experimental design and "control group" is associated with an experimental design

EXAMPLE OF AN EXPERIMENT

- Hypothesis: Eating in class improves test performance
  - Manipulated variable: Eating/not eating in class

- Step 1 (Pretest): Students take a quiz
- Step 2 (Intervention):
  - Participants are randomly assigned into the eating and then listen to a lecture
- Step 3 (Posttest): Students take another quiz

QUASI-EXPERIMENTS

- What is a quasi-experiment?
  - The prefix "quasi" means "almost," so a quasi-experiment is an "almost" experiment
  - Less controlled than a traditional experiment
    - NO random assignment / Don't always have a control group
    - Don't always include pretests and/or posttests
- Main reason for use? Feasibility

QUASI-EXPERIMENTAL AND EXPERIMENTAL DESIGNS

Quasi-Experimental	Experimental
• Can be challenging to identify a similar comparison group • Because program and comparison groups are different, results are considered less rigorous • Often less labor intensive and expensive than experimental design	• Most rigorous design option, so results tend to be more highly regarded • Often requires increased program recruitment • Applicant acceptance is randomly determined • Can be more labor intensive and expensive

QUASI-EXPERIMENTS FOR PROGRAM EVALUATION

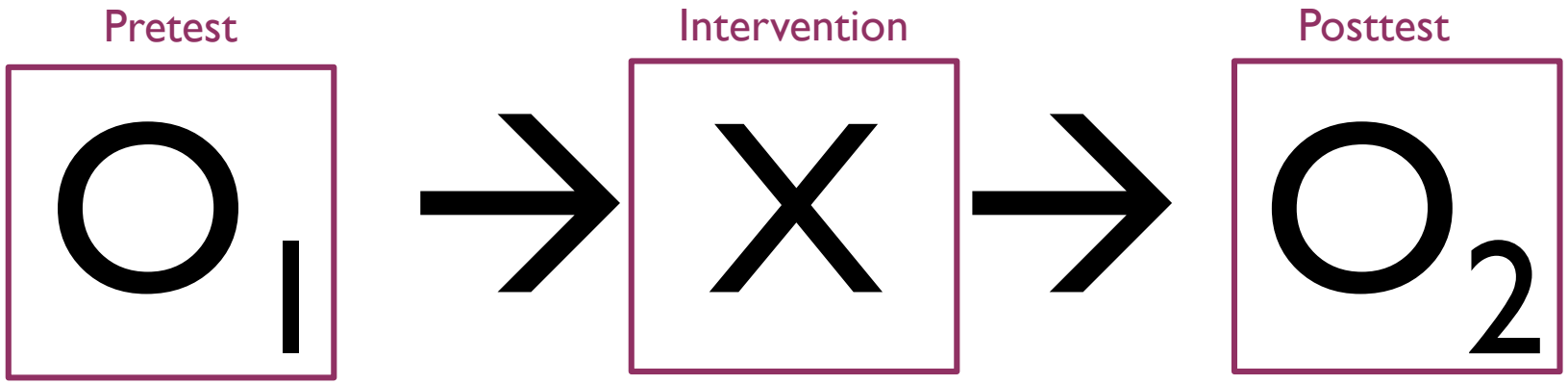
- Can be used to find evidence of changes in program participants
  - Ex: See if program leads to increased knowledge in participants
- Can be used to demonstrate program effectiveness
  - Ex: Provide evidence that the program itself was responsible for the increase in knowledge

TYPES OF QUASI-EXPERIMENTS

- One-group pretest-posttest
- Comparison group design
- Interrupted time series

ONE-GROUP PRETEST-POSTTEST DESIGN

- Compare the same group of people before and after an intervention
- If a change occurs, the intervention may be responsible



O = Observation
X = Intervention

COMPARISON GROUP DESIGN

- Two non-random groups are compared
  - One group receives the intervention; other serves as control group
  - Each group receives a different intervention
  - May or may not have pretests

Posttest-Only

Group 1: X → O

Group 2: O

Pretest-Posttest

Group 1: O → X → O

Group 2: O → → O

EXAMPLE OF COMPARISON GROUP DESIGN

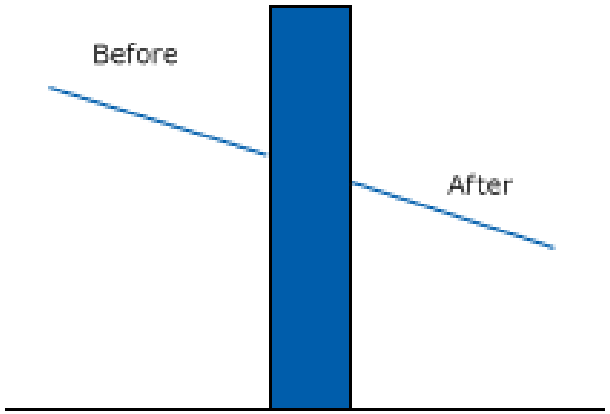
- Two groups of students take an knowledge test
- One group goes through the program; the other does not
- Both groups of students take a second knowledge test to check for differences

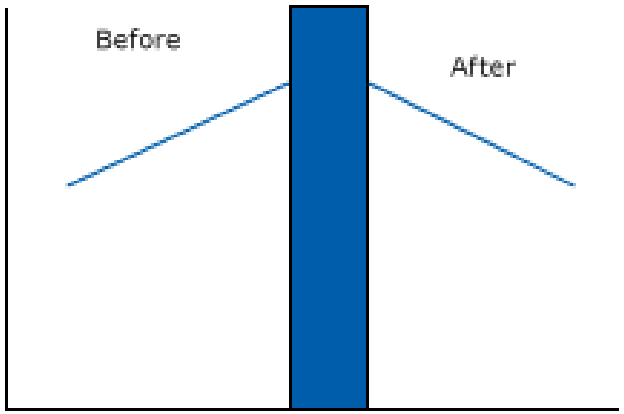
Group 1: O → X → O

Group 2: O → → O

INTERRUPTED TIME SERIES DESIGN

- One group is tested several times before and after an intervention; if a significant change in scores happens after the experiment, it can be attributed to the intervention
  - $O_1 \rightarrow O_2 \rightarrow O_3 \rightarrow X \rightarrow O_4 \rightarrow O_5 \rightarrow O_6$

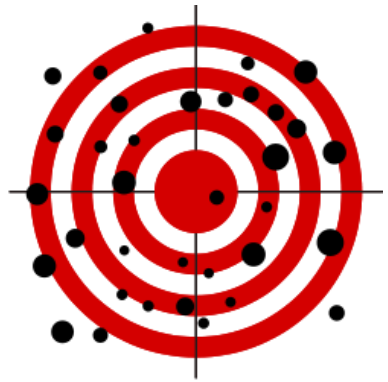




RELIABILITY AND VALIDITY

- **Reliability:** Consistency of results
- **Validity:** Accuracy of the results

Unreliable & Unvalid



Unreliable, But Valid

Reliable, Not Valid

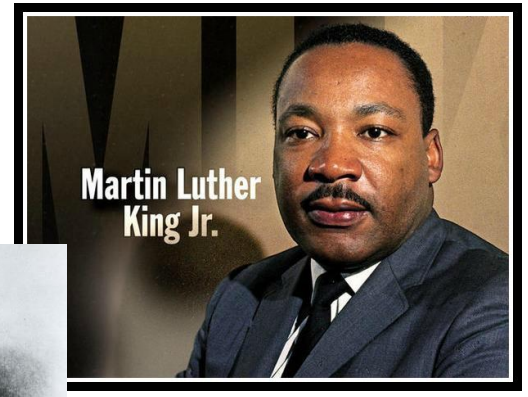
Both Reliable & Valid

THREATS TO VALIDITY

- History
- Maturation
- Testing
- Instrumentation
- Statistical Regression
- Selection Bias
- Experimental mortality

THREATS TO VALIDITY: HISTORY

- An outside event may occur during the course of the experiment which might influence results
  - Example: During a study on social inequality, a prominent civil rights leader is assassinated



THREATS TO VALIDITY: MATURATION

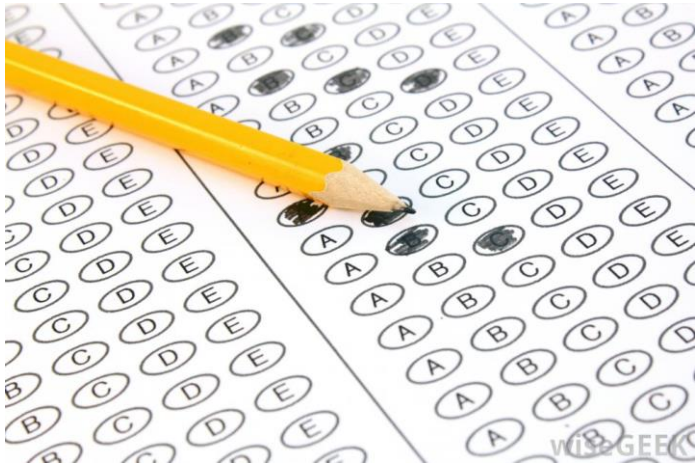
- People change and grow over the course of an experiment
- Can happen in both short- and long-term experiments
  - Example: Participants grow bored over the course of a study
  - Example: Students' improved reading scores are due to their natural progression

THREATS TO VALIDITY: TESTING

- The act of testing/experimenting itself influences performance
  - Example: By the time participants receive a posttest, they've become more sensitive to purpose of the study

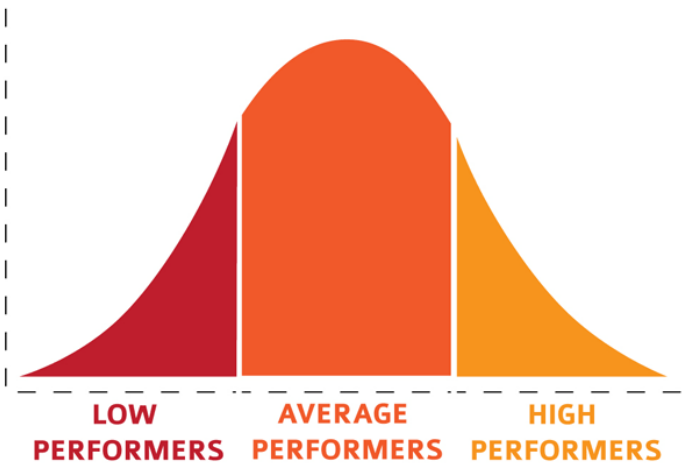
THREATS TO VALIDITY: INSTRUMENTATION

- If different measures are used in the pre and posttests, they may not be comparable
  - Example: Students score higher on a posttest because it was easier than the pretest, not because they learned more.



THREATS TO VALIDITY: STATISTICAL REGRESSION

- If participants start at an extreme, any changes may be falsely attributed to the experiment
- Also called "regression to the mean"
  - Example: In a study on students who have failed math classes, improvements are attributed to the intervention/program when in reality their scores improved because they had nowhere to go but up



THREATS TO VALIDITY: SELECTION BIAS

- The way experimental groups are selected leads to different results; not the experiment itself
  - Example: Students who opt-in to a study on extracurricular activities are inherently more motivated than those who do not

THREATS TO VALIDITY: EXPERIMENTAL MORTALITY

- Participants drop out of a study before it is completed, which may impact findings from the study
  - Example: The lowest performing students drop out of a tutoring study