



# Exploring Survey Methods

## ES 200: Introduction to Environmental Studies

Paul Manson  
Spring 2023

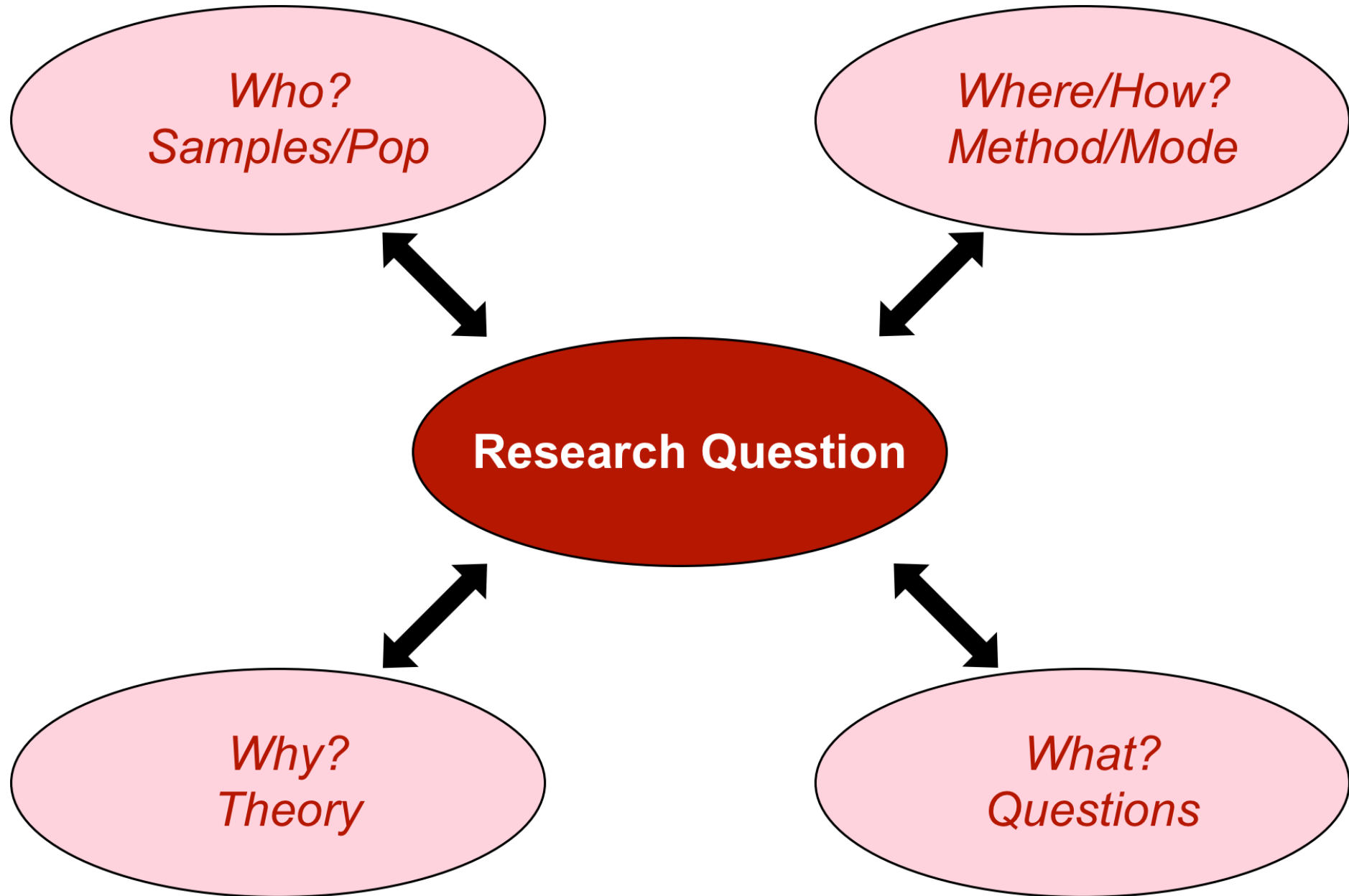
# Day One: Introduction to Survey Methodology



# Day One Outline

- 🕵️ Why: Motivations for Surveys and Theory
- 📞 Who: Sampling and Recruitment Challenges
- 💪 Where and How: Administering Surveys
- ✅ What: Basics in Survey Design
- 📊 Analyze Data





# What Are Surveys?

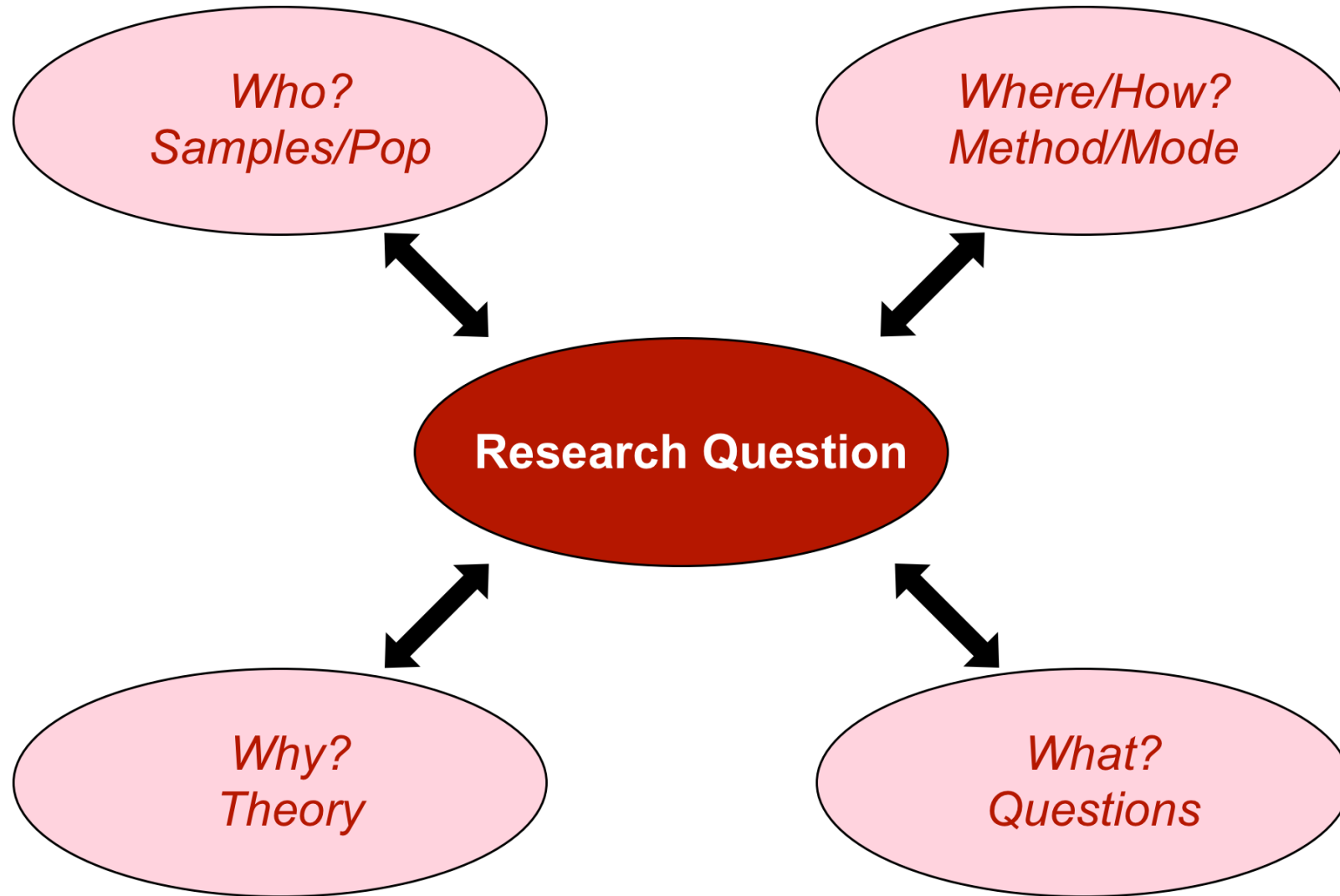
- Qualitative AND/OR Quantitative
- Many Types: Interviews, Questionnaires
- Respondent Focus - Who Completes Them
- Many Modes: In-person, Door-to-Door, Intercept, Mail, Phone, Online
- Allow for inferences from Sample to Population

## When Not to Use Surveys

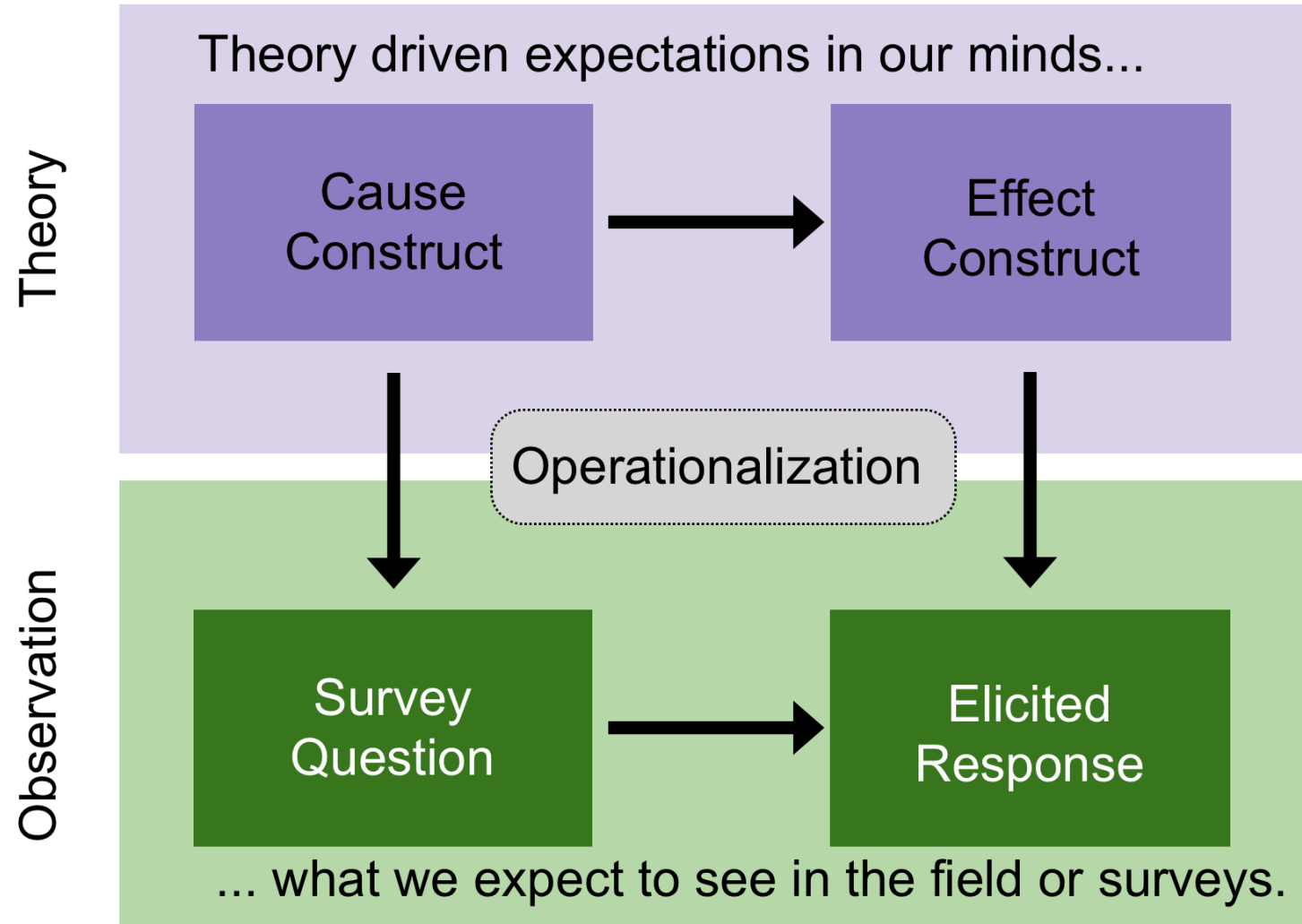
- Research Questions are Emergent
- Various Sources of Bias or Error Overwhelm
- Population Hard to Reach



# First: The Why



# The Why: Theory Rules Everything



# The Why: Developing Causal Relationships

🤔 What is a cause and effect you are interested in? What is the Theory?

- 🌍 Climate change beliefs are determined by party identification
- 🌲 Support for protecting natural areas is predicted by frequent visitation or use
- 🐳 Support for marine protected areas is driven by concerns over loss of marine health
- Other ideas?





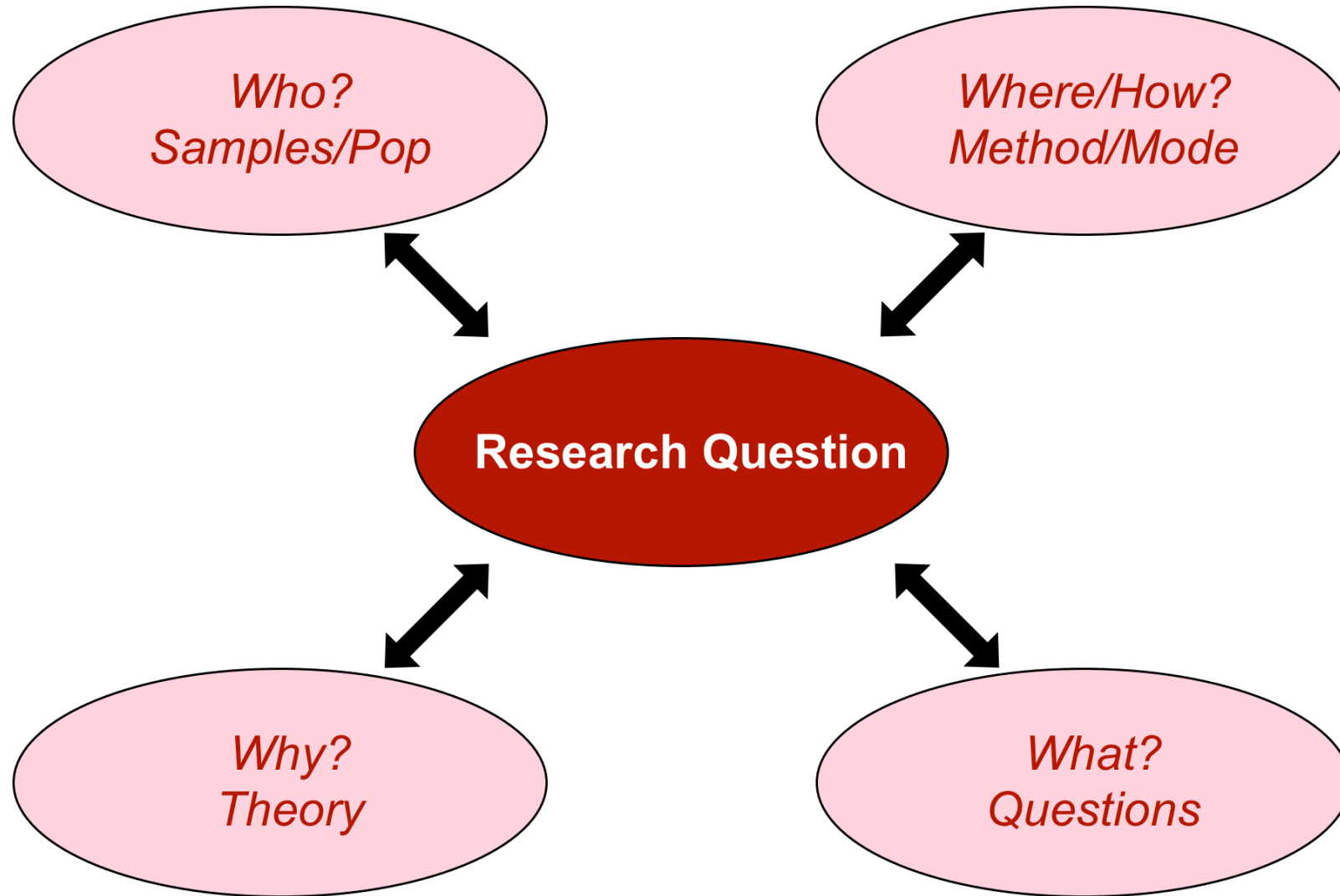
# The Why: Developing Causal Relationships

🤔 What would you ask to test these relationships?

- 🌍 Do you believe climate change is happening? ➡  
**Varies by party identification**
- 🌲 Willingness to pay for protecting open space, expressed in dollars. ➡  
**Varies by visit frequency**
- 🐳 Hypothetical vote for expanding marine protected areas ➡  
**Varies by perception on marine health**
- Other ideas?



## Second: The Who



# The Who: Getting People to Participate

We *could* just ask everyone! (aka Census)

- Time intensive
- Expensive
- Might actually be less accurate! (Might...)

## Sampling Approaches

- Random
  - Pure Random, Stratified, Systematic
- Purposive and Quota
- Intercept
- Snowball
- Convenience



# Sample Method Drives Inferences

- Random sample allows for inference from sample to population
- Samples need to be carefully drawn:
  - **Population:** (Universe) Definition: Who do we want to Represent?
  - **Frame:** How we Find our Sample
  - **Sample:** Drawn from Frame, assuming proper *coverage*

## Remember Unit of Analysis

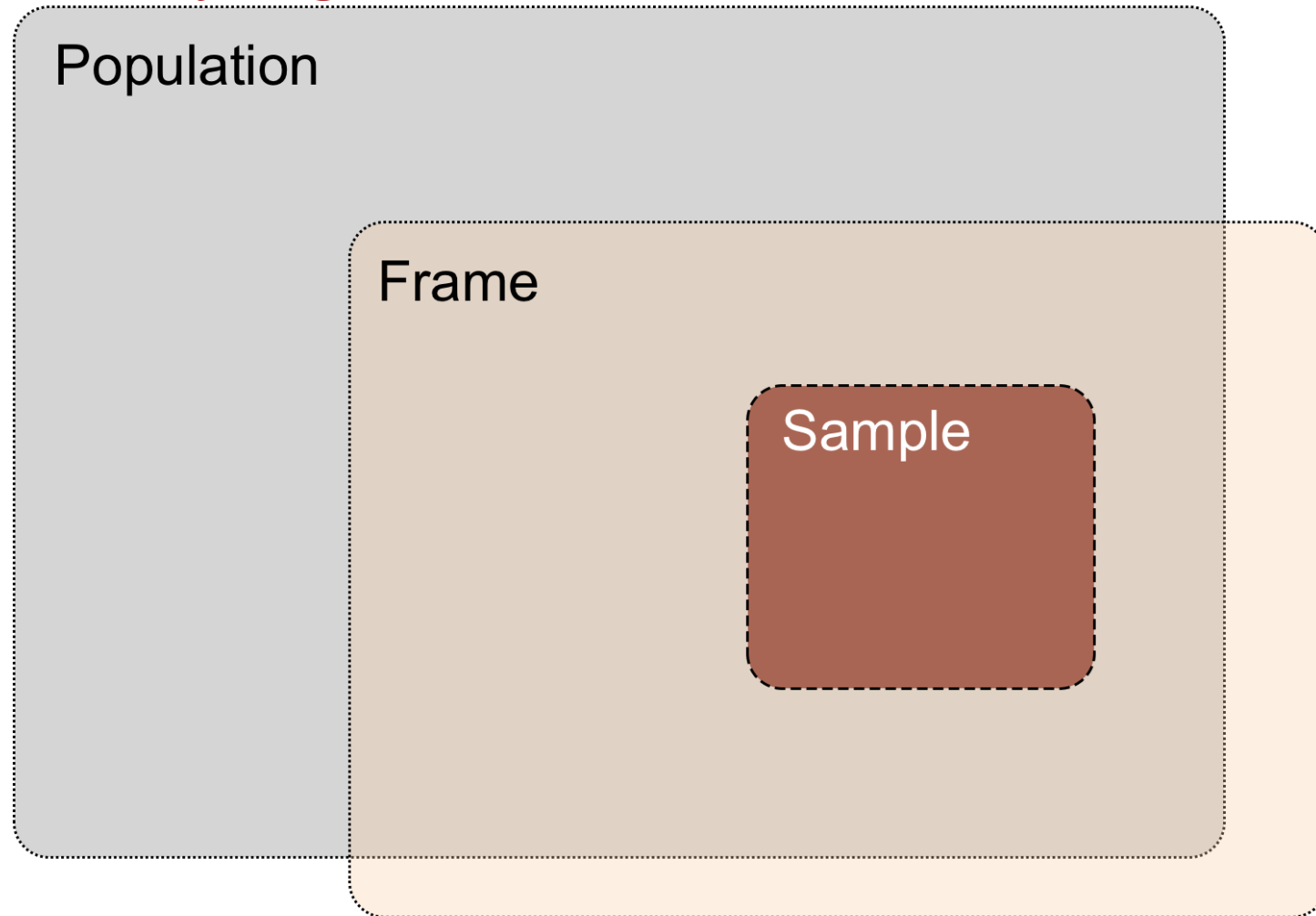
- Is it the individual? The household? A firm or organization? -

## Sources of Error

- **Coverage Error:** The Frame has a mismatch!
- **Sampling Error:** Size, Response Rates



# The Who: Sampling and Recruitment



# How Many Surveys Do we Need?

95% Conf Level	10% MOE		5% MOE		3% MOE	
Population Size	50-50	80-20	50-50	80-20	50-50	80-20
100	49	38	80	71	92	87
200	65	47	132	111	169	155
400	78	53	196	153	291	253
1000	88	58	278	198	517	406
2000	92	60	322	219	696	509
10000	95	61	370	240	965	640
40000	96	61	381	244	1040	672
100000	96	61	383	245	1056	679
1000000	96	61	384	246	1067	683

From Dillman et al 2009



# Confidence Intervals vs Levels

- **Confidence Intervals** or two times the margin of error, measures the how sure we are of the answer given.
- **Confidence Level** measures how sure we are that the sample, if drawn again, would have the same result.

**Example:** If we are working with a sample size of 384, for a population of over 10 million people, we have a 5% margin of error using the 95% confidence level.

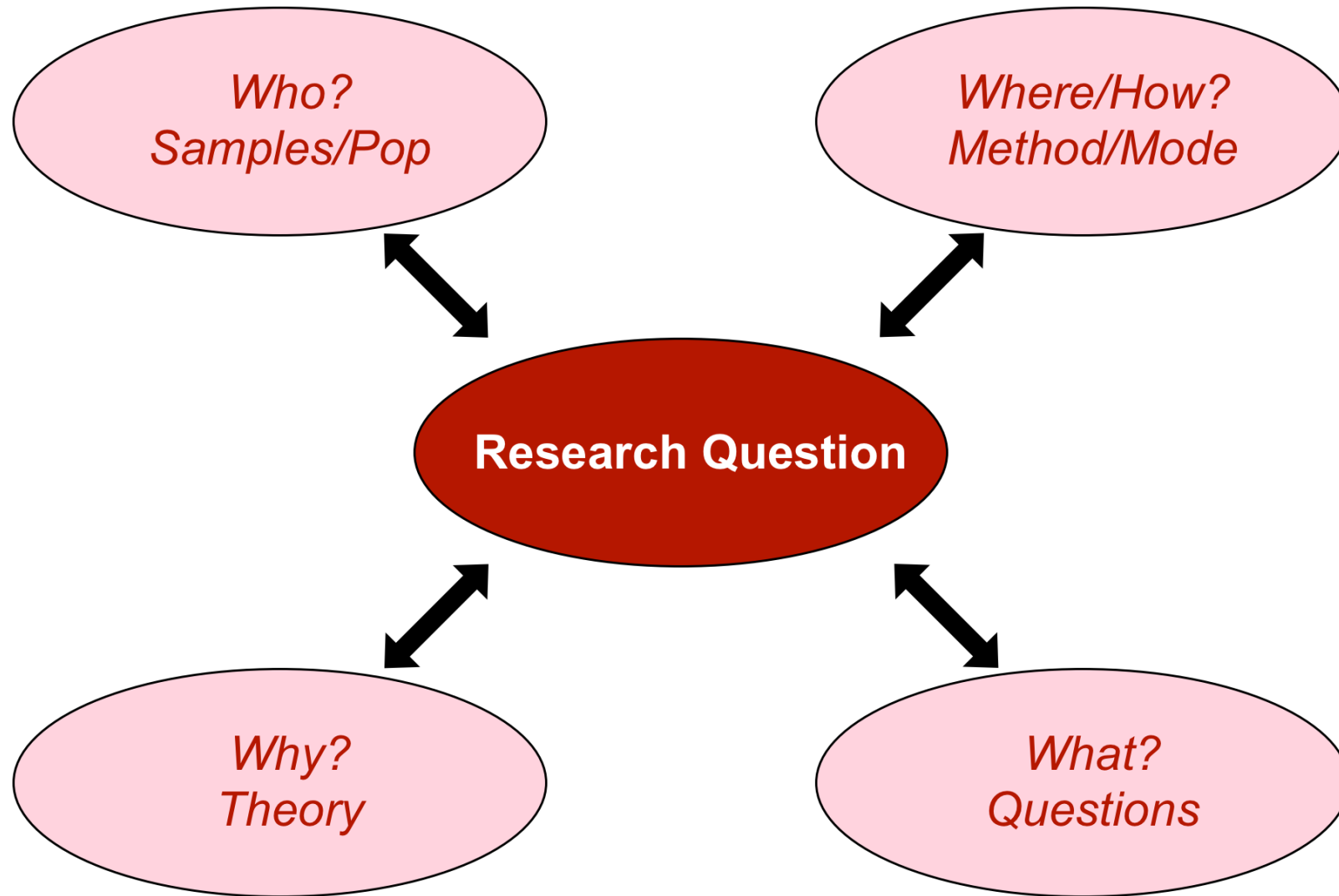
That means our confidence interval is plus or minus 5% on the sample statistic.

If 52% of the population said they were very sure climate change was happening, we would say we are **95% sure** that the true percentage is **between 47% and 57%**.

If we wanted to narrow that range with the 3% margin of error, we'd **need 1067 responses** (not 384!)



## Next: The Where and How





# How: The Many Modes of Survey Administration

## How we connect to Respondents

- Online
- Mail
- Phone
- Text/SMS
- Web Panel
- Door to Door
- Intercept

Each of these involve some form of **recruitment** and may use multiple ways!

**Example:** Post card sent to household inviting to online survey. Follow up with paper survey for non-responses. Final phone call to explore response non-bias.



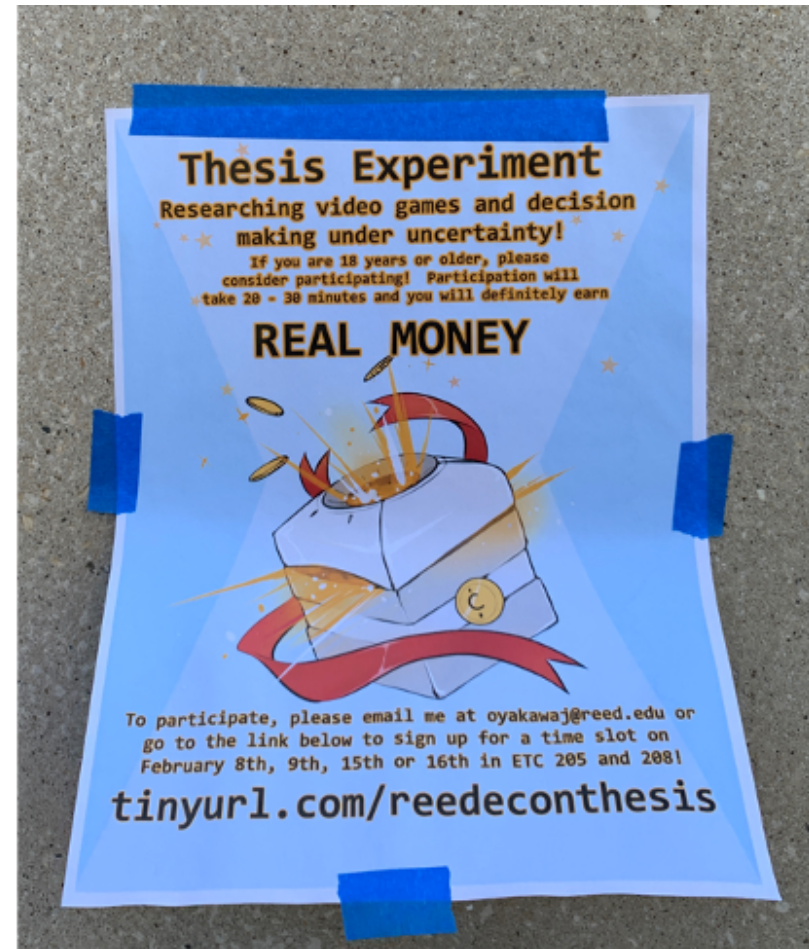
# How: The Many Modes of Survey Administration

## How we encourage Responses

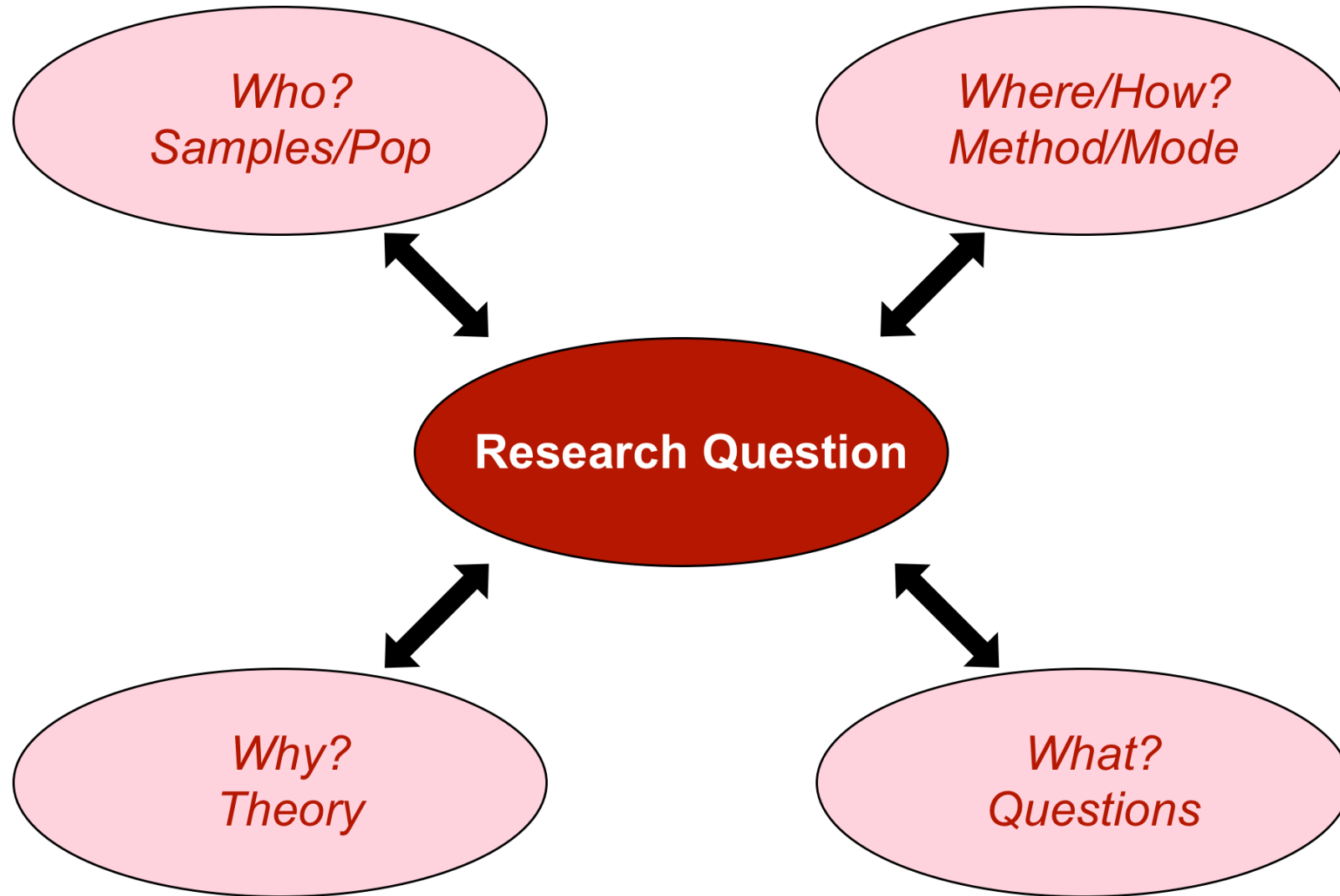
- Make a Connection: Cover Letters and Design
- Partnerships: Trusted Entities
- Incentives
- Use your "Student Card"
- Consider Timing - Know your Population

## Response Rates

- Depends on: Non-Contact, Eligibility, Partial Completion, Full Completion



# Finally: The What



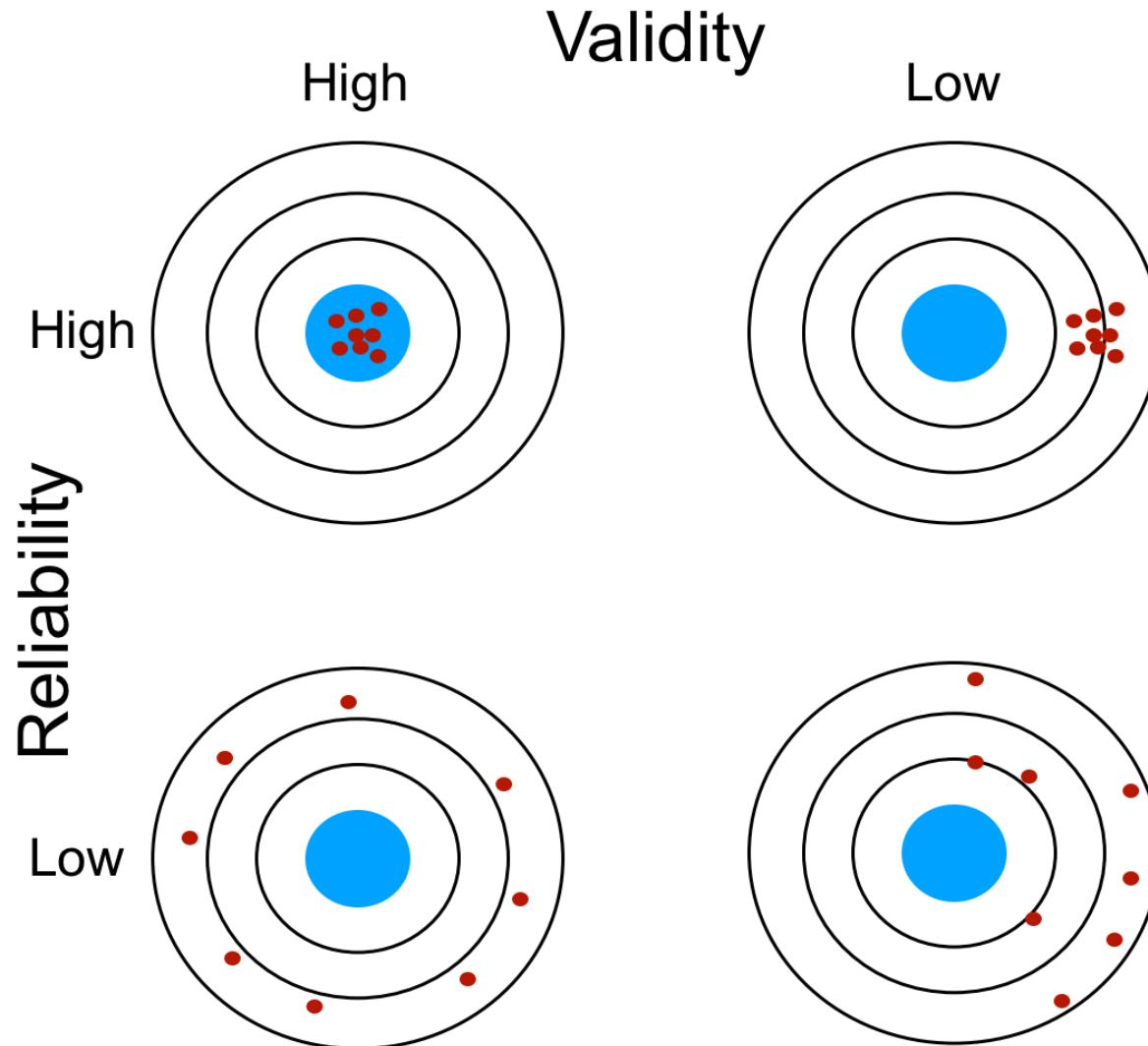
# The Best for the Last: Questions and How to Create Them

- Measurement Challenges
- Best Practices
- Things to Avoid
- Connect Back to Theory

🤔 What is a Survey Question you Might Ask?



# Measurement Challenges

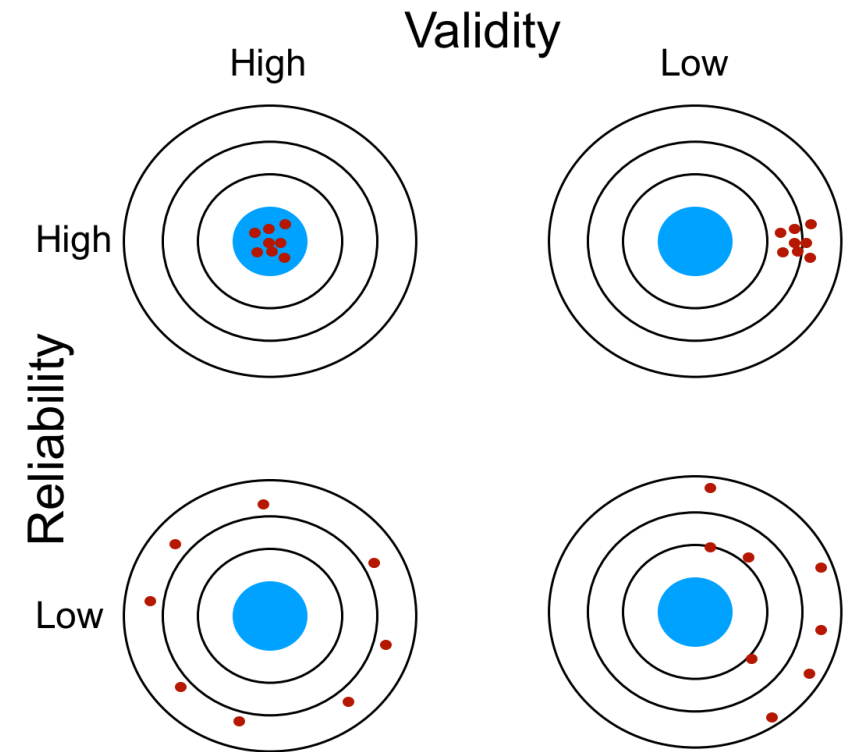


# Measurement Challenges

- Are we measuring what we want to measure (**Validity**)?
- Are we successful when repeatedly using the measure (**Reliability**)?

## Some examples

- High Validity and High Reliability:  
**Respondent's age**
- High Validity and Low Reliability:  
**Assess classroom inclusion and climate**
- Low Validity and High Reliability:  
**Are you racist?**
- Low Validity and Low Reliability:  
**Do you support strenological expansion?**



# Types of Questions

- Open Ended Questions: Text entry or "in their own words"
- Closed-Ended: Choices are pre-determined
- Other
  - Image Based
  - Participatory GIS
  - Text Interactions



# Choices on Question Design

## Open-Ended

- Information rich (Good and Bad!)
- Does not prejudge right response
- Allows for uncertainty, where researcher does not know what to expect
- A **lot** of work to analyze

## Close-Ended

- Researcher can control responses (analysis strength)
- Comparisons are easier
- Quicker for respondent
- You have to be very careful on choices you offer!





# Types of Close-Ended Questions

- Dichotomous
  - Yes or No, Checkbox Questions
- Categorical
  - Often demographic or descriptive
- Ordinal (or acting like it!)
  - Likert scale questions: Strongly Agree to Strongly Disagree
- Ratio (rare)
  - Money, thermometers, scales with sliders



# Anatomy of a Question (aka Item)

- Introduction
- Stem
- Response Options
  - Scale
  - Anchors
  - Odd vs Even Scales (5-pt or 4-pt)
  - Don't Know or Refusal Options
- Presentation Order
  - Questions can be randomized, or reversed
  - Scale can be randomized, or reversed
- Skip or Presentation Logic
  - Present questions based on other responses (aka Fork)
- Soft Requires and Hard Requires
  - Soft: Respondent offered a second chance if they skip
  - Hard: Respondent can not proceed without answer (consent example)



# Danger! Things to Avoid!

- Double Barrel Questions:  
"Do you agree Renn Fayre is safe and fun?"
- Biased Questions:  
"Do you support climate heroes working to stop oil companies?"
- Overly Leading:  
"Do you love Reed College?"
- Ambiguous or Jargon-loaded:  
"Do you never oppose fracking on BLM lands?"
- Mis-matched Stem and Response Options:  
"Do you recycle at home? Every week, occsasisionally, never."
- Unbalanced Response Options:  
"Disagree, Neither Agree nor Disasgree, Agree, Stronly Agree"

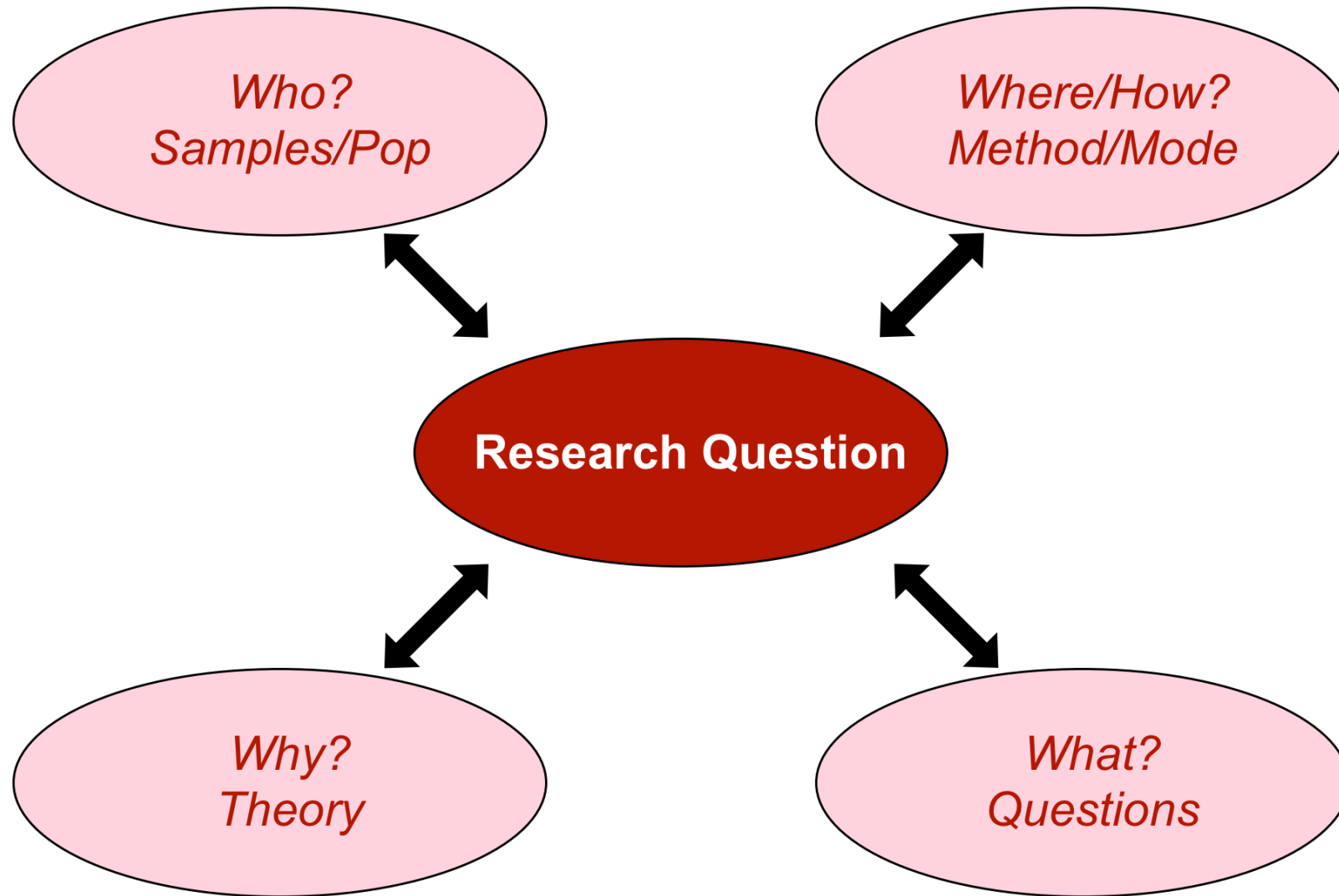


# Awesome! Things to Definately Do! ❤️

- Use Visual Design  
White space, approachable design or colors, images.
- Language and Accessibility:  
Can everyone in your population participate? Translation or other tools to help with accessibility?
- Make the Benefit Clear:  
How can your research change the world?
- Incentives: Small tokens go a long way. Can be raffle or payment for all.
- Do no Harm:  
Remember your ethical obligations, do no harm to respondents



You Made It! 🏆



# Preview on Analysis

## Response Rates and Weights

- Even if we draw the perfect sample, we miss some people.
- Use of weights to adjust results to capture over and under-representation

## Types of Weights

- **Design Weights:** We might add more people we know are likely to get missed into our sample draw.
- **Post-Stratification/Non-Response:** Once we have our sample, we can adjust to match population characteristics.



# Discuss Survey Codebook

 What questions interested you?

 What other variables would you use to understand question responses?

 In groups: Discuss questions and analyses you might explore! (Keep these ideas, we will use them!)



# Day Two: Analyzing Survey Data





# Day Two Outline

- 🚚 Accessing Existing Survey Datasets
- 🧺 Loading into R and Wrangling
- 🧭 Exploring the Data
- 📊 Analyze Data
- 🏗️ Designing in Qualtrics

