Presenter: Dennis Wang

Survey Research

From *Ways of Knowing in HCI (Olson & Kellogg, 2014)*

Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge

Outline

What is Survey Research (In the context of HCI)?

What Constitutes Good Work?

What is Survey?

What is Survey?

- Method of gathering information by asking questions to a subset of people, the results of which can be generalized to the wider target population

What is Survey?

- Survey in Ancient time
  - General decision & policy making
    - Since ancient times, societies have measured their populations via **censuses**.
    - Political polls amplified public interest in surveys
  - Social psychologists aimed to minimize questionnaire biases and optimize data collection

What is Survey?

- Survey in HCI
  - HCI before Internet
    - Understand users' experiences with computer hardware and software
    - *Computer-collected survey responses* with those *from a printed questionnaire*
  - Web-based Survey Research (HCI after Internet)
    - In 1994, the Georgia Institute of Technology started annual online surveys to understand Internet usage and users and to explore Web-based survey research
    - Online applications widely adopted surveys to measure users' satisfaction, unaddressed needs, and problems experienced, in addition to user profiling

What Questions Can be Answered by Survey?

Provide insights into users' attitudes, experiences, intents, demographics, and psychographic characteristics

When Are Surveys Appropriate ?

When Should We Avoid Using a Survey ?

When Surveys Are Appropriate

Surveys are appropriate when needing to **represent an entire population**

- **Differences** between groups of people
- Identify **changes over time** in people's attitudes and experiences

Attitudes

Benchmark attitudes toward an application or an experience

Measure customer satisfaction with online banking immediately

Intent

"Why did you visit this website?"

Task Success

reliably quantify levels of success

When Surveys Are Appropriate

More Examples!

User characteristics

Collect users' *demographic* information, *technographic* details

Interactions with technology

Self-report on social, psychological, and demographic variables

Awareness

Understanding people's awareness of existing technologies

When to Avoid Using a Survey

Precise behaviors

Users will struggle to recall their exact sequence of actions

Gathering this information from *log data*, if available, will always be more accurate

Underlying motivations

Unable to explain why they take certain actions or prefer one thing over another

Ethnography or *Contextual inquiry*

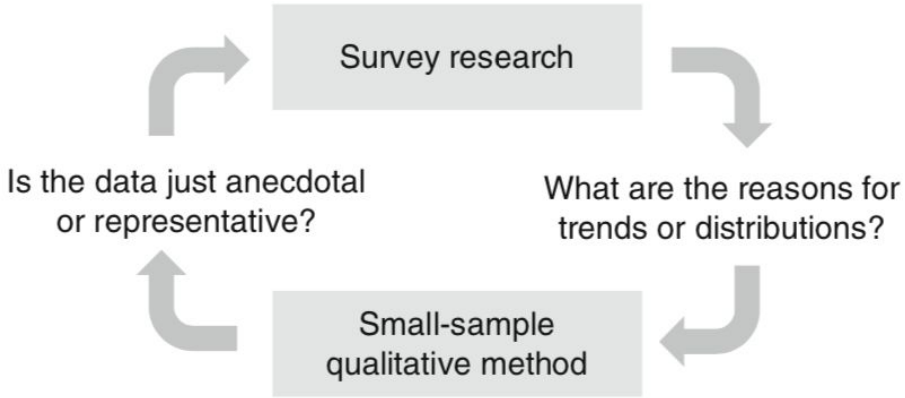
Usability evaluations

Can measure task success but may not explain why people cannot use a particular application

Task-based observational research and *Interview methods*

Using Surveys with Other Methods

- **Qualitative (e.g. Usability Study)**
  - follow previous qualitative studies to help quantify specific observations
  - first to identify the range of frustrations or goals, followed by qualitative interviews and observational research



Using Surveys with Other Methods

- **Log data (Psychophysiological data)**
  - Log data may show that one experimental version drives more traffic or engagement, but the survey may show that users were less satisfied or unable to complete a task
  - log data can further validate insights from a previously conducted survey
  - Psychophysiological data (e.g. EEG)
- **A/B Test**
  - satisfaction and self-reported task success can be measured and analyzed in parallel with behavioral differences observed in log data

What Constitutes Good Work

What Constitutes Good Work

1. Research goals and constructs
2. Population and sampling
3. Questionnaire design and biases
4. Review and survey pretesting
5. Implementation and launch
6. Data analysis and reporting

Research goals and constructs

- **Research Goal** (e.g. *overall satisfaction in Happiness*)
- **Constructs**
  - Unidimensional attributes that cannot be observed
  - identified constructs should then be converted into one or multiple survey questions
  - identified from prior primary research or literature reviews
  - e.g. “overall satisfaction,” “perceived speed,” and “perceived utility”
- **Survey Questions**
  - When the survey-appropriate research goals have been identified, they should be matched to constructs
    - evaluate its unique relationship with the higher level goal
    - cognitive pretesting can be used to determine whether respondents are interpreting the constructs as intended by the researcher

Research goals and constructs

- **Determine if the survey approach is appropriate after research goals and constructs are defined**
  - Survey constructs focus on results which will directly address research goals and inform stakeholders' decision making (E.g. ensure not excessive on nice-to-know questions)
  - Temporality of result
    - longitudinal comparisons or for one-time?
  - Number of responses
    - provide the appropriate level of precision

Population and sampling

- Determining who and how many people to survey.
- Sampling frame
  - because reaching everyone in the population (i.e., a census) is typically *impossible* and *unnecessary*
  - Reachable: the set of people who the researcher is able to contact for the survey
  - Identical: perfect sampling frame is identical to the population (but often a surveys' sampling frame is only a portion of the population)

Population and sampling

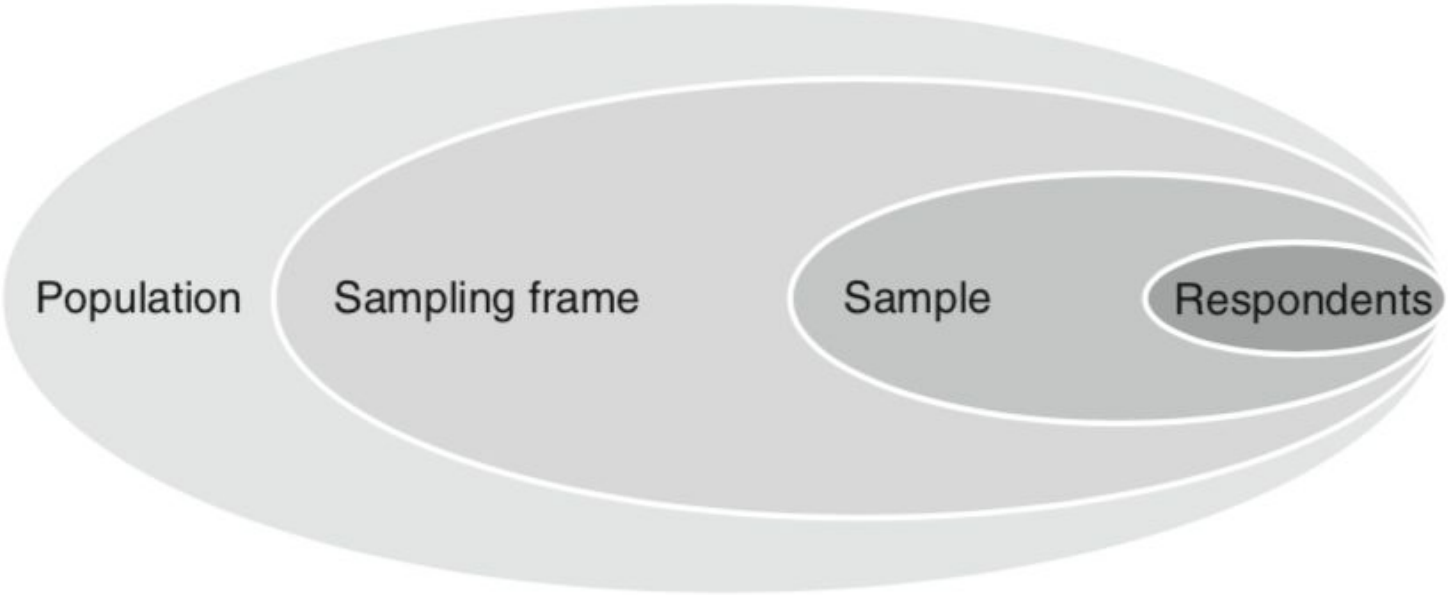


Fig. 3 The relationship between population, sampling frame, sample, and respondents

Population and sampling

- Understanding the satisfaction of an application’s users
  - **Population**:

Population and sampling

- Understanding the satisfaction of an application’s users
  - **Population**: everyone that uses the application

Population and sampling

- Understanding the satisfaction of an application's users
  - **Population**: everyone that uses the application
  - **Sampling frame**:

Population and sampling

- Understanding the satisfaction of an application's users
  - **Population**: everyone that uses the application
  - **Sampling frame**: users that are actually reachable
    - (exclude those who have abandoned the application, anonymous users, and users who have not opted in to being contacted for research)

Population and sampling

- Understanding the satisfaction of an application's users
  - **Population**: everyone that uses the application
  - **Sampling frame**: users that are actually reachable
    - (exclude those who have abandoned the application, anonymous users, and users who have not opted in to being contacted for research)
  - **Sample**:

Population and sampling

- Understanding the satisfaction of an application's users
  - **Population**: everyone that uses the application
  - **Sampling frame**: users that are actually reachable
    - (exclude those who have abandoned the application, anonymous users, and users who have not opted in to being contacted for research)
  - **Sample**: people from the sampling frame who are invited to take the survey

Population and sampling

- Understanding the satisfaction of an application's users
  - **Population**: everyone that uses the application
  - **Sampling frame**: users that are actually reachable
    - (exclude those who have abandoned the application, anonymous users, and users who have not opted in to being contacted for research)
  - **Sample**: people from the sampling frame who are invited to take the survey
  - **Respondent**: people who actually finish the survey

Population and sampling

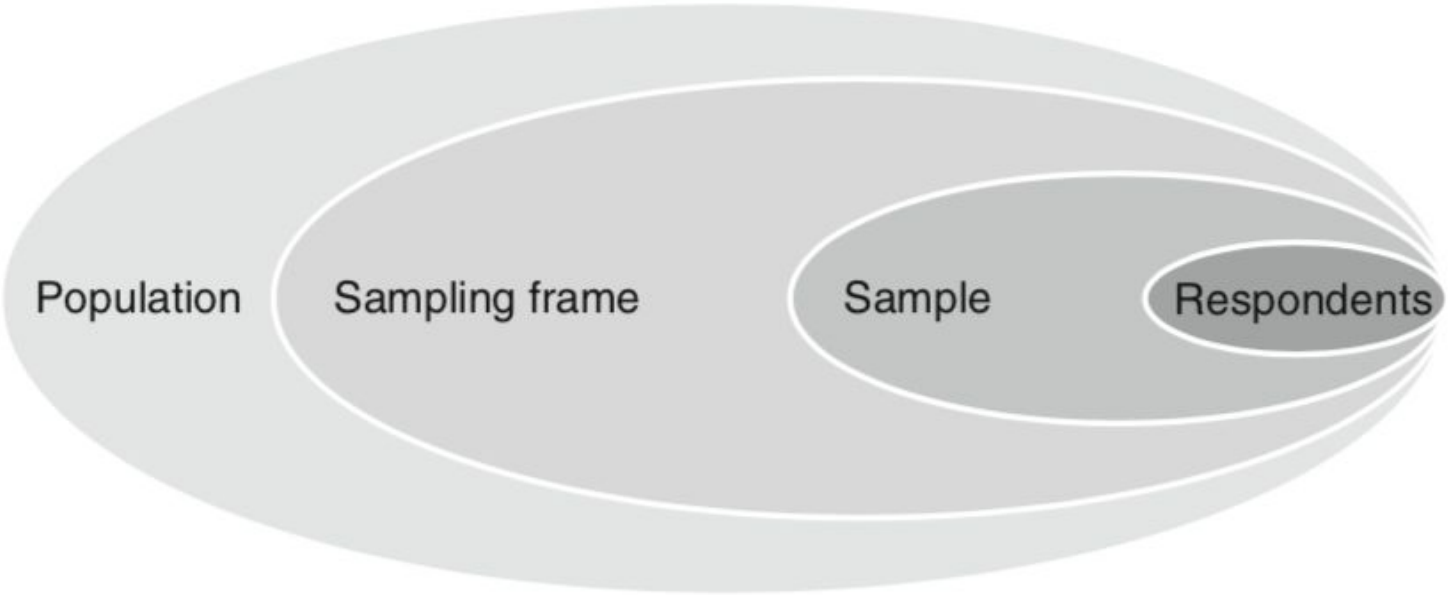


Fig. 3 The relationship between population, sampling frame, sample, and respondents

Population and sampling

- *Sampling* a population in Sampling frame
- ***Probability***:
  - every person in the sampling frame has an equal, non-zero chance of being chosen for the sample
  - Minimize sample-bias: the sample is selected completely randomly
  - random digit telephone dialing / address-based mail surveys / list-based samples for e-mail invitations
  - pop-up surveys or in-product links

Population and sampling

- ***Non-probability***:
  - While probability sampling is ideal, it is often impossible to reach and randomly select from the entire target population, especially when targeting small populations
  - self-selected surveys / snowball recruiting / convenience samples
  - high sampling bias? comparing key characteristics of the target population with those from the actual sample
  - e.g. recruiting Down's syndrome patients through special interest mailing lists

Population and sampling

- Determining the Appropriate Sample Size
  - carefully determine the target sample size for the survey, i.e., the number of survey responses needed.
  - Krejcie and Morgan's formula (1970)

Confidence level	90%				95%				99%			
Size of population / Margin of error	10%	5%	3%	1%	10%	5%	3%	1%	10%	5%	3%	1%
10	9	10	10	10	9	10	10	10	9	10	10	10
100	41	73	88	99	49	80	92	99	63	87	95	99
1000	63	213	429	871	88	278	516	906	142	399	648	943
10,000	67	263	699	4035	95	370	964	4899	163	622	1556	6239
100,000	68	270	746	6335	96	383	1056	8762	166	659	1810	14227
1,000,000	68	270	751	6718	96	384	1066	9512	166	663	1840	16317
100,000,000	68	271	752	6763	96	384	1067	9594	166	663	1843	16560

Population and sampling

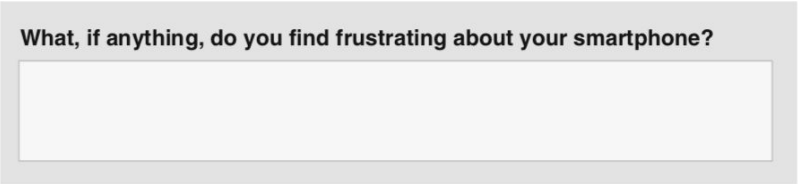
- Mode and Methods of Survey Invitation
  - Mail or written surveys, phone surveys, in-person surveys, and Internet surveys
  - Internet Survey
    - **Advantage**: Easy access to large geographic regions / Cost savings during survey invitation / Ability to customize the questionnaire to specific respondent groups using skip logic
    - **Disadvantage**: *coverage error* (a potential mismatch between the target population and the sampling frame)
      - Similar issue happened in *Crowdsourcing in HCI Research*

Questionnaire design and biases

- Why we need to carefully design the questionnaire
  - Measurement error
    - the deviation of the respondents' answers from their true values on the measure
  - Couper (2000), measurement error in self-administered surveys can arise from the respondent (e.g., lack of motivation, comprehension problems, deliberate distortion) or from the instrument (e.g., poor wording or design, technical flaws).

Questionnaire design and biases

- Types of Survey Question
  - Open-ended questions
    - Universe of possible answers is unknown or Too many options
    - Measuring quantities with natural metrics (age, length, or frequency)
    - unable to access Information from log data
    - Measuring qualitative aspects of a user's experience
  - Closed-ended questions
    - Universe of possible answers is known and small enough to be easily provided
    - Rating a single object on a dimension
    - Measuring quantities without natural metrics (importance, certainty, or degree)



Questionnaire design and biases

- Questionnaire Biases
  - check the phrasing of each question for potential biases that may bias the responses
  - Satisficing
    - use a suboptimal amount of cognitive effort to answer questions
    - Cognitive ability / Motivation to answer is low.
      - Question difficulty is high at one of the four stages
    - How to avoid: Avoid Complexity

Questionnaire design and biases

- Questionnaire Biases
  - Acquiescence Bias
    - concur with the statement independent of its substance (e.g. tend to agree with things)
    - When Respondent's Cognitive ability is low / Question difficulty is high
    - How to avoid: Avoid questions with agree/disagree, or similar answer options. Ask construct-specific questions

Questionnaire design and biases

- Questionnaire Biases
  - Social Desirability
    - answer questions in a manner they feel will be positively perceived by others
    - People who not inclined to go against the social norm
    - How to avoid: allowed to answer anonymously or the survey should be self-administered

Questionnaire design and biases

- Questionnaire Biases
  - Response Order Bias
    - tendency to select the items toward the beginning (i.e., primacy effect) or the end (i.e., recency effect)
    - Strongest when the list of answer options is long
    - How to avoid: switch order
  - Question Order Bias
    - Order effects also apply to the order of the questions in surveys
  - Other Types of Questions to Avoid

So far...

- Research goals and constructs
- Population and sampling
- Questionnaire design and biases

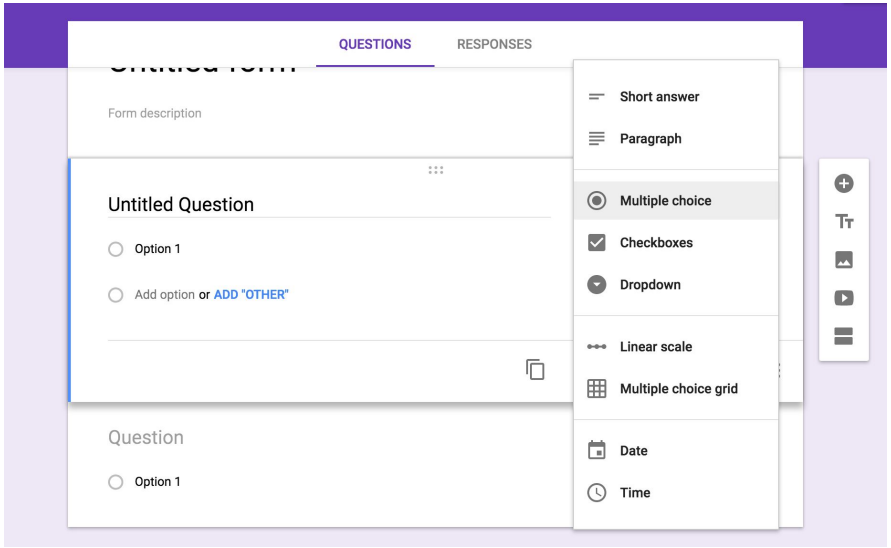
Questions so far?

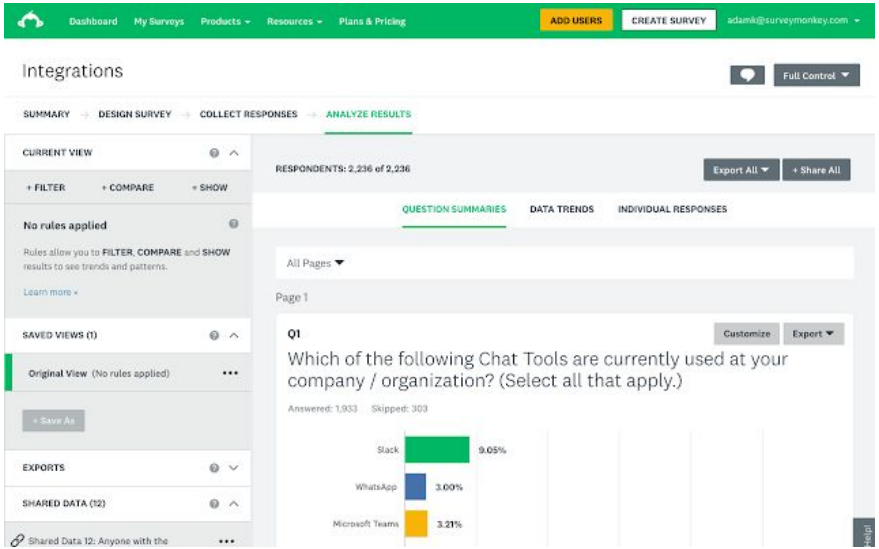
Review and survey pretesting

- have potential respondents take and evaluate the survey in order to identify any remaining points of confusion
- Cognitive Pretesting
  - a small set of potential respondents is invited to participate in an in-person interview where they are asked to take the survey while using the think-aloud protocol
- Field Testing
  - Piloting the survey with a small subset of the sample will help provide insights that cognitive pretests alone cannot
  - Metrics: Look for *break-off points* and *long completion times*

Implementation and launch

- Tools / Platform for Survey
  - ConfirmIt, Google Forms, Kinesis, LimeSurvey, SurveyGizmo, SurveyMonkey, UserZoom, Wufoo, Zoomerang, and so on...





Implementation and launch

- Piping Behavioral Data into Surveys
  - support the ability to combine survey responses with other log data
  - accurately understand the relationship between respondent characteristics and their behaviors or attitudes
  - find that certain types of users or the level of usage may correlate with higher reported satisfaction
- Monitoring Survey Paradata
  - With the survey's launch, researchers should monitor the initial responses as well as survey paradata to identify potential mistakes in the survey design.
  - Click-through rate / Completion rate / Response rate / Break-off rate / Completion time

Implementation and launch

- Maximizing Response Rates
  - In order to gather enough responses to represent the target population with the desired level of precision, response rates should be maximized.
  - “Total Design Method” (Dillman, 1978)
    - Four mailings: the initial request with the survey on week one, a reminder postcard on week two, a replacement survey to non-respondents on week four, and a second replacement survey to non-respondents by certified mail on week seven.
    - Leveraging Social Exchange Theory, & personalizing the letters
    - Later did an experiment on Internet-based Tests
  - Providing an incentive

Data analysis and reporting

- Preparing and exploring the data
- Thoroughly analyzing the data
- Synthesizing insights for the target audience of this research

Data analysis and reporting

- Preparing and exploring the data
  - Data Preparation and Cleaning
    - Cleaning and preparing survey data before conducting a thorough analysis are essential to identify low-quality responses that may otherwise skew the results
    - What are some principles for data cleaning?
      - Duplicate responses / Speeders / Straight-liners and other questionable patterns / Missing data and break-offs
      - Question-by-question: Low inter-item reliability / Outliers / Inadequate open-ended responses

Data analysis and reporting

- Thoroughly analyzing the data
  - Closed-Ended Responses
    - descriptive statistics are fundamental.
    - Quantitative analysis
  - Open-Ended Comments
    - holistic understanding of the phenomena
    - Coding
      - Deductive / Inductive / Intra-rater reliability

Data analysis and reporting

- Synthesizing insights for the target audience of this research
  - Assessing Representativeness
    - the degree to which the results accurately represent the target population
    - understanding non-response
  - Reporting Survey Findings
    - result + detailed description of the survey methodology

What Constitutes Good Work

1. **Research goals and constructs**
    a. Research Goal -> Constructs -> Survey Qs
2. Population and sampling
3. Questionnaire design and biases
4. Review and survey pretesting
5. Implementation and launch
6. Data analysis and reporting

What Constitutes Good Work

1. Research goals and constructs
   a. Research Goal -> Constructs -> Survey Qs
2. **Population and sampling**
   a. Sample Frame
   b. Sampling
3. Questionnaire design and biases
4. Review and survey pretesting
5. Implementation and launch
6. Data analysis and reporting

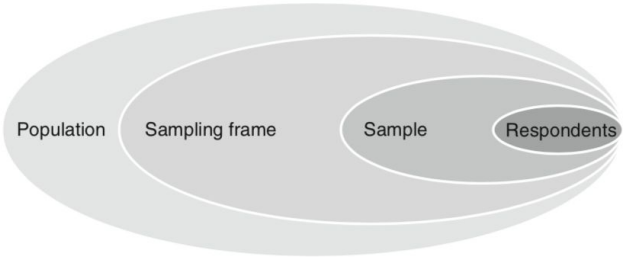


Fig. 3 The relationship between population, sampling frame, sample, and respondents

What Constitutes Good Work

1. Research goals and constructs
   a. Research Goal -> Constructs -> Survey Qs
2. Population and sampling
   a. Sample Frame
   b. Sampling
3. **Questionnaire design and biases**
   a. Measurement error & Questionnaire bias
4. Review and survey pretesting
5. Implementation and launch
6. Data analysis and reporting

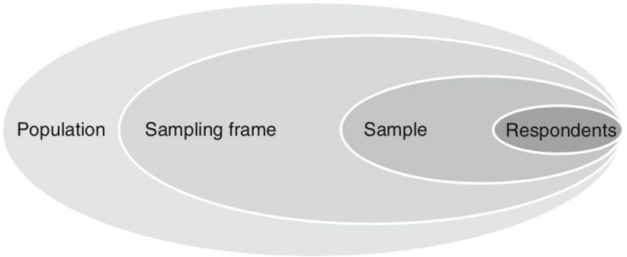


Fig. 3 The relationship between population, sampling frame, sample, and respondents

What Constitutes Good Work

1. Research goals and constructs
   a. Research Goal -> Constructs -> Survey Qs
2. Population and sampling
   a. Sample Frame
   b. Sampling
3. Questionnaire design and biases
   a. Measurement error & Questionnaire bias
4. **Review and survey pretesting**
   a. Pilot Study
5. Implementation and launch
6. Data analysis and reporting

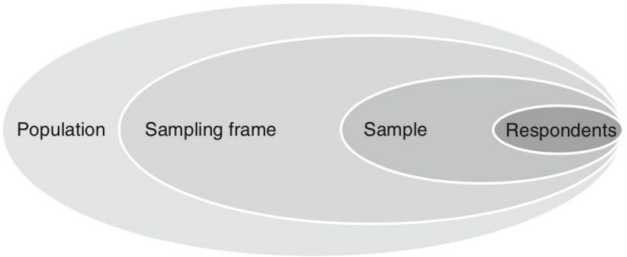


Fig. 3 The relationship between population, sampling frame, sample, and respondents

What Constitutes Good Work

1. Research goals and constructs
   a. Research Goal -> Constructs -> Survey Qs
2. Population and sampling
   a. Sample Frame
   b. Sampling
3. Questionnaire design and biases
   a. Measurement error & Questionnaire bias
4. Review and survey pretesting
   a. Pilot Study
5. **Implementation and launch**
6. Data analysis and reporting

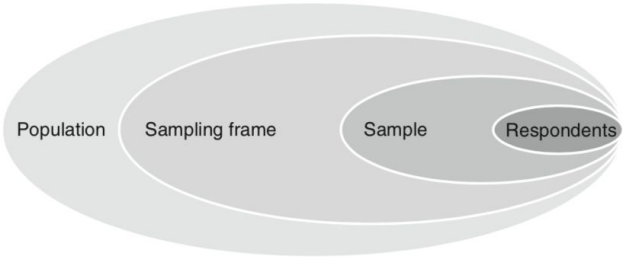


Fig. 3 The relationship between population, sampling frame, sample, and respondents

What Constitutes Good Work

1. **Research goals and constructs**
    a. Research Goal -> Constructs -> Survey Qs
2. Population and sampling
    a. Sample Frame
    b. Sampling
3. Questionnaire design and biases
    a. Measurement error & Questionnaire bias
4. Review and survey pretesting
    a. Pilot Study
5. Implementation and launch
6. **Data analysis and reporting**

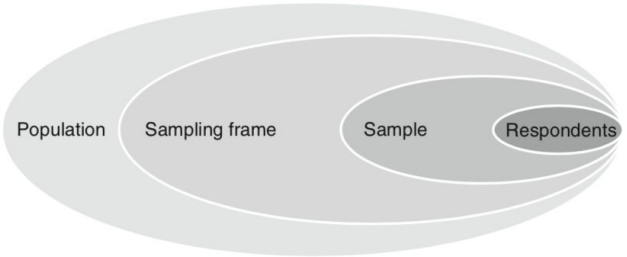


Fig. 3 The relationship between population, sampling frame, sample, and respondents

Questions?