

Survey Design: Samples and Questionnaires

APEC SKILLS WORKSHOP

OCTOBER 23, 2024

CAROLINE KRAFFT, HUMPHREY SCHOOL OF PUBLIC AFFAIRS, UNIVERSITY OF MINNESOTA

Overview



Do you need to do a survey?



Sampling for surveys



Data collection management



Design of questionnaires and question types



Data collection digital tools

Do you need to do a survey?

3

- ▶ Surveys are complicated, time consuming, and expensive
 - ▶ Especially to do well
- ▶ Before you invest in survey data collection, investigate whether there are (publicly available) surveys or other data sources that already have the data you need to answer your research question
- ▶ Consider whether you might be able to add a module or questions to someone else's survey, rather than run a whole new survey
- ▶ If you cannot get data by other means, then can consider developing a survey
- ▶ Additionally, if you are running a randomized controlled trial, you will (usually) need to run survey(s) to estimate impact

Example: Graduate tracer surveys

Graduate tracer surveys are designed to capture critical labor market information on the graduates of a specific institution

There is not administrative data nor are there existing surveys that capture this information

I have supported Egyptian universities in developing, collecting, and analyzing data from graduate tracer surveys

The data from tracer surveys can inform both improvements within the university and the career advising provided to students

When to survey?

5

- ▶ If using descriptive or quasi-experimental methods, may be able to do a single cross-sectional survey
 - ▶ Example: Regression discontinuity design based on age cutoff for pre-primary, examining impact on maternal outcomes
- ▶ Some descriptive and quasi-experimental methods may require longitudinal data
 - ▶ Labor market transitions
 - ▶ (Time) fixed effects, difference-in-difference
 - ▶ May be possible to capture this in one survey retrospectively, but beware recall issues
- ▶ For randomized controlled trial (RCT, experimental) methods, typically undertake:
 - ▶ Baseline survey, (intervention), (maybe) midline survey, endline
 - ▶ Occasionally can skip baseline

When to conduct a graduate tracer survey?

6

- Can collect data at multiple points in time for the same individuals (longitudinal data)



Education program



Graduating cohort
Exit survey
Contact information



Graduate Tracer Survey
Graduation + 1 year



Graduate Tracer Survey
Graduation + 3 years



Sampling & longitudinal data collection

Sampling breakdown

Who do you want to generalize to?

- The (theoretical) population

What population can you get access to?

- The study population

How can you get access to them

- The sampling frame

Who is in your study?

- The sample

Sample frame & sample



A sample is a smaller (but trying for representative) collection of units from the population, used to determine information about that population



Why sample?

Resources (time, money) and workload

Gives results with known accuracy that can be calculated mathematically (standard errors)



The sampling frame is the list from which the potential respondents are drawn



For tracer studies, the (theoretical) population is a graduating cohort from an institution

Example: "Class of 2024" from Cairo University

Sample representativeness

10

- ▶ Goal: a representative sample
- ▶ Need to consider your unit of observation
 - ▶ Households? Individuals? Firms?
- ▶ Three factors that influence sample representativeness
 - ▶ Sampling procedure
 - ▶ Sample size
 - ▶ Participation (response rate)
- ▶ When might you sample the entire population?
 - ▶ When your population is small
 - ▶ When you have extensive resources
 - ▶ When you don't expect a very high response rate

Longitudinal data collection

11

A **longitudinal study** is a research design that involves repeated observations of the same subjects (and variables) over time

Tracer studies are one type of longitudinal study which sample a cohort (a group who experienced a common event (graduation) in a selected period (a particular year) and perform observations at intervals through time (graduation, +1 year, +3 year).

Example: Tracer study sample frames 12

Ideal sample frame: All the students graduating in a particular year (cohort)

- Requires up-to-date contact information for each student
 - Ideally, home address, phone, and (non-institutional) email, possibly social media handles (e.g. LinkedIn)
 - Ideally, alternate contacts (parents' phone #s, for example)
- This ideal sample frame (if successfully sampled) allows you to generalize to all the students of that cohort
- Requires university having and sharing this contact information
- A university-wide exit survey can be valuable for the university, and as a source of contact information

More readily available sample frame: University career center contacts

- Can gather contact information for each individual who contacts career center (meeting, event, etc.)
- Can only generalize (make claims about) students who worked with career center, not all students
 - Students who contact career center may be very proactive and do better-than-average on the labor market
 - Students who contact career center may be those struggling to find jobs and do worse-than-average on the labor market

Sampling methods

13

- ▶ **A simple random sample** is one where every item in the population has the same chance of being sampled
 - ▶ Example: Take entire list of the graduating cohort, using statistical software, randomly sample 10%
- ▶ **Strata** are identifiable sub-groups
 - ▶ Singular of strata: stratum
 - ▶ Examples: different specializations, different degrees (BA vs. MA), gender
 - ▶ These strata need to have data in the sample frame to sample by
 - ▶ When there are clear strata, can use **stratified random sampling**: random sample within each stratum
 - ▶ Simplest: Proportional to stratum's size in population
 - ▶ Complex: Over-sample small groups (e.g. small specializations) to ensure a big enough sample size. Then **must** weight the sample by the inverse of the sampling rate for all statistics.
- ▶ **Cluster sampling** randomly selects clusters
 - ▶ **Clusters** are geographic or other naturally occurring groups
 - ▶ Often referred to as "primary sampling units" (PSUs)
 - ▶ Example: To get contact information, could do an exit survey with a *sample* of classes and the graduating students in those classes

Sample Size Determination



Determining the sample size for a study is a crucial component of study to include sufficient numbers of subjects so that statistically significant results can be detected.



"How large a sample do I need?"



The answer will depend on the aims, nature and scope of the study and on the expected result. All of which should be carefully considered at the planning stage.

Sampling Error

- ▶ A measurement of a random sample has the best chance (probability) of approximating the expected value of the population.
- ▶ Sample measurements will reflect some random sampling error.
- ▶ Random sampling error simply refers to the fact that the sample measure is not the likely to be the same as the population, and thus, there is a certain amount of error associated with samples (difference between the sample measurement and the actual population value).
- ▶ Random sampling error is expected and O.K.
 - ▶ Statistical procedures, including sample size estimation, take sampling error into account.
- ▶ Bias, or sampling bias, is NOT O.K.
 - ▶ Can lead the researcher to draw inaccurate conclusions about the population.

Margin of error & confidence level

16

- ▶ The **margin of error** expresses the amount of sampling error in the survey, i.e., the range of uncertainty around an estimate of a population measure, assuming normally distributed data.
 - ▶ For example, if 60 % of the sample claims to use a computer in their work, a 5 % margin of error would mean that actually 55–65 % of the population use computers.
- ▶ The **confidence level** indicates how likely the reported metric falls within the margin of error if the study were repeated.
 - ▶ A 95 % confidence level, for example, would mean that 95 % of the time, observations from repeated sampling will fall within the interval defined by the margin of error. Commonly used confidence levels are 99, 95, and 90 %.

Sample Size (in Cells), Population Size, Margin of Error and Confidence Level

- ▶ The sample size you need depends on the population size, the margin of error you are willing to accept for key measures, and your confidence level
 - ▶ Example: 370 observation sample required to have a 5% margin of error at a 95% confidence level for a population of 10,000

- ▶ R or online calculators can do sample size calculations
 - ▶ <https://www.qualtrics.com/blog/calculating-sample-size/>

- ▶ Keep in mind
 - ▶ This is the realized sample size – non-response will require initially sampling more individuals
 - ▶ That some outcomes, such as wages, will only be available for sub-samples and so larger samples overall may be required to achieve the same level of precision

Confidence level	90%				95%				99%			
Margin of error Size of population	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

Power calculations

18

- ▶ If you are undertaking an RCT, need to do power calculations to determine your sample size
- ▶ Figure out a meaningful minimum detectable effect size (MDE)
 - ▶ Based on literature or descriptive data
- ▶ Crunch (power, clustersampsi commands in Stata) based on:
 - ▶ Means, standard deviations for outcome
 - ▶ Number of treatment arms
 - ▶ Clusters (if applicable) and intra-cluster correlation
 - ▶ Alpha (sig. level, false positive), Power (false negative, prob. falsely failing to reject null)

Sample size tradeoffs

- ▶ If sample (“n”) is
 - ▶ Large
 - ▶ Increase accuracy
 - ▶ Costly and complex
 - ▶ Small
 - ▶ Decrease accuracy
 - ▶ Less costly

Data collection management

Managing contact information

21

- ▶ Contact information may be received in a variety of formats
 - ▶ Written, email, online forms, over the phone
- ▶ Need to unify into a single, standardized database in order to sample
- ▶ Need to ensure the data, which are personally identifiable information, are securely stored
 - ▶ Note: Need to **get IRB approval** of your sample, questionnaire, study well before starting research!
- ▶ Need a process for updating contact information in the database
 - ▶ Students are quite likely to move soon after graduation, and may change phone #s, emails, etc.
 - ▶ Will need to collect updated contact information for any longitudinal studies (observing individuals +1 and +3 years after graduation)

Non-response: Problems

22

- ▶ When individuals in the sample do not respond to the survey, it creates a number of problems:
 - ▶ Smaller sample size
 - ▶ Non-response bias due to selective non-response
 - ▶ Example: graduates who obtained jobs are busy and less likely to respond
- ▶ Sources of non-response:
 - ▶ Failure to contact
 - ▶ Examples: bad phone, busy signal, # or email
 - ▶ Refusal
 - ▶ Need to ask for consent, individuals may refuse
- ▶ Calculate and report the response rate: $(\text{Respondents/sample}) * 100$

Non-response: Solutions

23

- ▶ Reaching respondents:
 - ▶ Ensure multiple, up to date forms of contact information
 - ▶ Reach out through multiple forms of contact (email, text, call)
- ▶ Contacted but not responding:
 - ▶ Ensure the invitation to participate is appealing
 - ▶ Send (friendly) reminders
 - ▶ Different times, days
- ▶ If reached:
 - ▶ Provide an incentive for completion
 - ▶ Keep your surveys short and simple
 - ▶ Avoid sensitive information or let individuals opt out
 - ▶ For example, do not ask if they are paying income taxes
 - ▶ Allow for reporting income continuously, categorically, or not at all
 - ▶ Set expectations accurately at start of survey (tell participants how long it will take)

Attrition in longitudinal studies

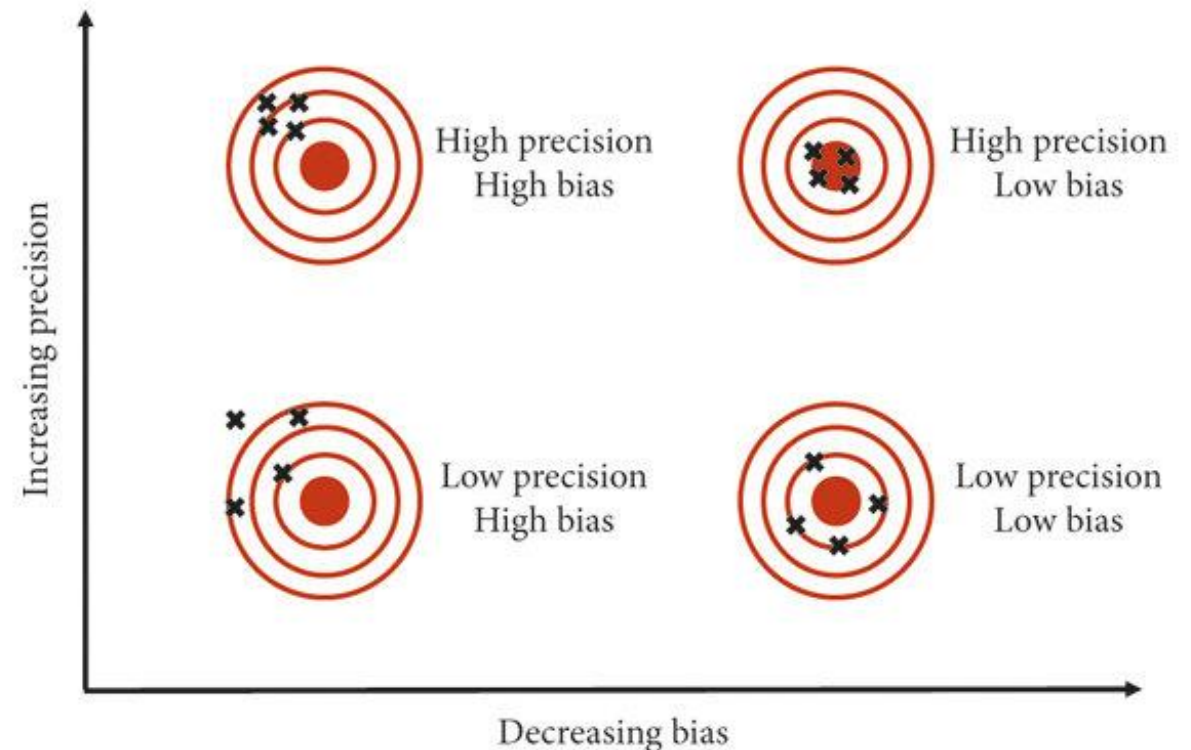
24

- ▶ **Attrition** occurs when participants in one round of the study do not respond in the next round
 - ▶ Example: Nahla responds to the exit survey and +1 year tracer survey, but not the +3 year tracer survey
- ▶ **Attrition bias** occurs because there tend to be systematic differences between the people who do not respond and those who continue
 - ▶ Similar problem to non-response bias, can compound initial non-response bias
 - ▶ Someone who obtained a job between +1 and +3 may then attrite at +3 year survey
 - ▶ In RCTs, attrition related to treatment huge concern
- ▶ Consequences:
 - ▶ Reduces sample size (develop initial sample to ensure a large enough sample even with attrition)
 - ▶ Increased bias
- ▶ Solutions:
 - ▶ Same measures as addressing non-response bias
 - ▶ Collecting up-to-date contact information at the end of previous surveys
 - ▶ Mechanisms (e.g. alumni office?) for updating contact information

Tradeoffs: precision and bias

25

- ▶ Different data collection methods are going to have potential tradeoffs between sample size, response rate, and therefore precision and bias
 - ▶ Prioritize sample with the least bias that can achieve an adequate sample size
 - ▶ A much larger sample size is not useful if the resulting sample is biased
 - ▶ Example:
 - ▶ If have emails and phone #s for everyone:
 - ▶ A phone survey is likely to have a higher response rate and more representative sample than an emailed survey,
 - ▶ But requires staff time/budget for collecting the survey data, likely requiring a smaller sample
 - ▶ Estimating, for example, earnings for four different groups, better off with a smaller (less precise) but unbiased sample than a precise but biased sample



Questionnaire design and question types

Qualities of a good questionnaire

27



Meets the research objectives

Can you answer your research questions?

Can be helpful to develop a pre-analysis plan, or draft a report template with all the tables and headers (but empty) and ensure all the data you want is captured in the questionnaire



Encourage complete and accurate responses

Make sure respondents fully understand questions



Easy for respondents to give the information and researchers to analyze the responses

No responses that are paragraphs of text



Brief and interesting

Better off with a good response rate and accurate responses for a short questionnaire than a low and biased response to a long questionnaire

Steps of questionnaire design

Decide what information is required

Make a rough listing of the questions

Refine the question phrasing

Develop the responses/response format

Plan sequencing and skip patterns

Program questionnaire

Pre-test the questionnaire and revise

What information is required?

29

Fundamental tracer study questions

- ▶ What happens to graduates after leaving the education/training institution?
- ▶ Were they able to get paid employment in an acceptable time?
- ▶ Do they use the skills and knowledge they have gained in the education/training institution? If not, what are the reasons?
- ▶ What are the skills and competencies demanded in the labor market?

Main objectives

- ▶ To measure the employability of graduates
 - ▶ Employment situation;
 - ▶ Time to get the first employment;
 - ▶ Duration of job search;
 - ▶ Salary/income;
 - ▶ Position;
 - ▶ Economic sector;
 - ▶ Main work tasks/duties (occupation)
 - ▶ Working conditions (time, benefits);
 - ▶ Job satisfaction
 - ▶ Use of competencies and required competencies;
 - ▶ Usefulness of study/training program;
- ▶ To collect feedback from graduates to improve the university
 - ▶ Further education and training (education pathways)
 - ▶ Need for further education and training;
 - ▶ Evaluation of the study/training program;
 - ▶ Strengths and weaknesses of the study/
 - ▶ Training program;
 - ▶ Proposals for improvements.

Rough listing of questions

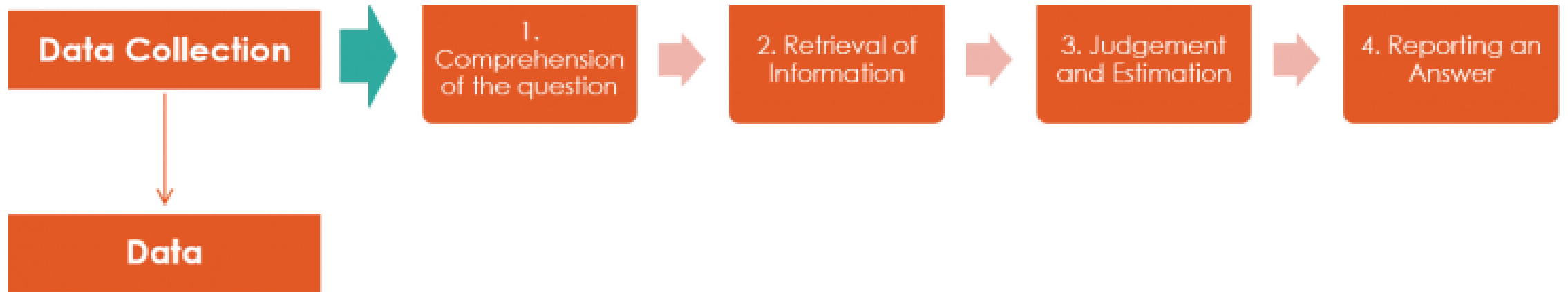
30

- ▶ Write out the questions you might want to ask respondents (don't worry too much about the wording yet)
- ▶ May require multiple questions to get at the key information of interest
 - ▶ "Unemployment" is best measured by a combination of questions on whether the individual is working, wants to work, is available to work, etc.
 - ▶ May want to ask start dates and end dates rather than duration
- ▶ Check for redundancy
 - ▶ Do not need to ask for date of birth and age
 - ▶ Do not need to ask for data already have administratively or from previous waves
- ▶ Example:
 - ▶ Demographics
 - ▶ Sex
 - ▶ Date of birth
 - ▶ Marital status
 - ▶ Labor market status
 - ▶ Employed in past 7 days
 - ▶ Attached but temporarily absent
 - ▶ If not employed, anything to search past 4 weeks yes/no
 - ▶ If not searching, why not
 - ▶ Job characteristics (if working)
 - ▶ Employment status
 - ▶ Occupation
 - ▶ Economic Activity
 - ▶ Sector

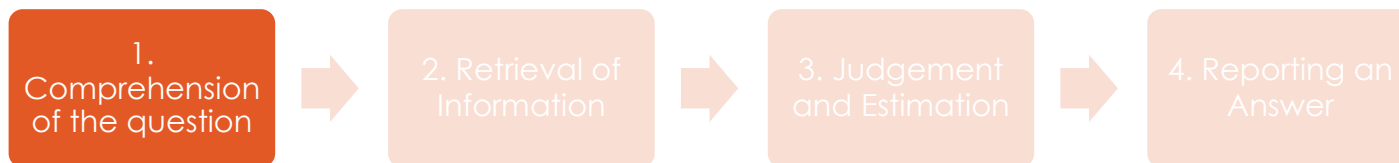
Question wording

31

- ▶ Don't "reinvent the wheel!"
 - ▶ Existing questionnaires, question banks can provide standardized, tested wording
- ▶ Each question should be SMART:
 - *Specific*: Asks one precise question
 - *Measurable*: Quantifies accurate and unbiased information
 - *Accessible* (understandable): Easy to comprehend and clearly defined
 - *Relevant*: Measures a key or intermediate outcome
 - *Time-bound*: Has a clear specification of the relevant time period
- ▶ Goal: minimize measurement error
 - ▶ Measurement error is the difference between a measured quantity and its true value



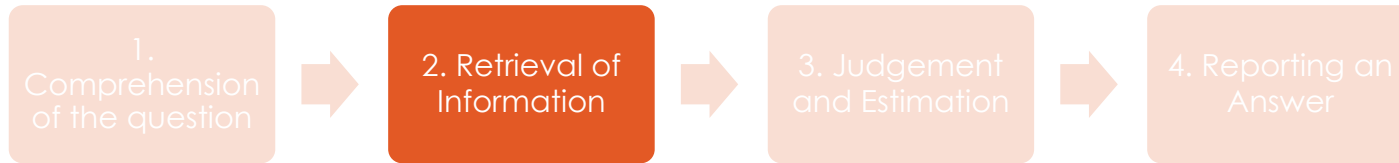
Process of answering a question



Step 1: Comprehension

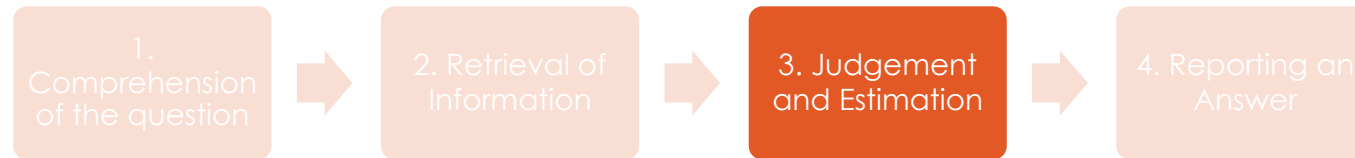
1.1 Total monthly income, before taxes





Step 2: Retrieval

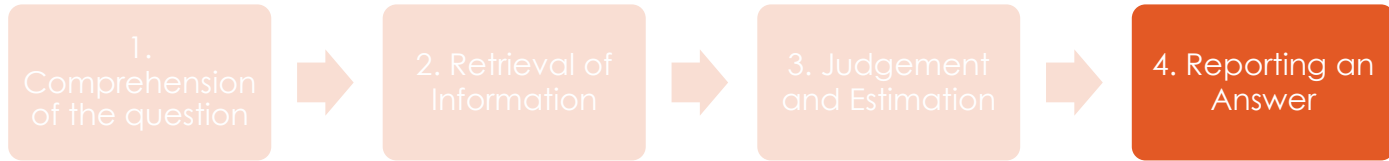




Step 3: Estimation/Judgement

Social = \$ 200 per month
Workers' Compensation = 0
Pension = \$220 per month
What else??

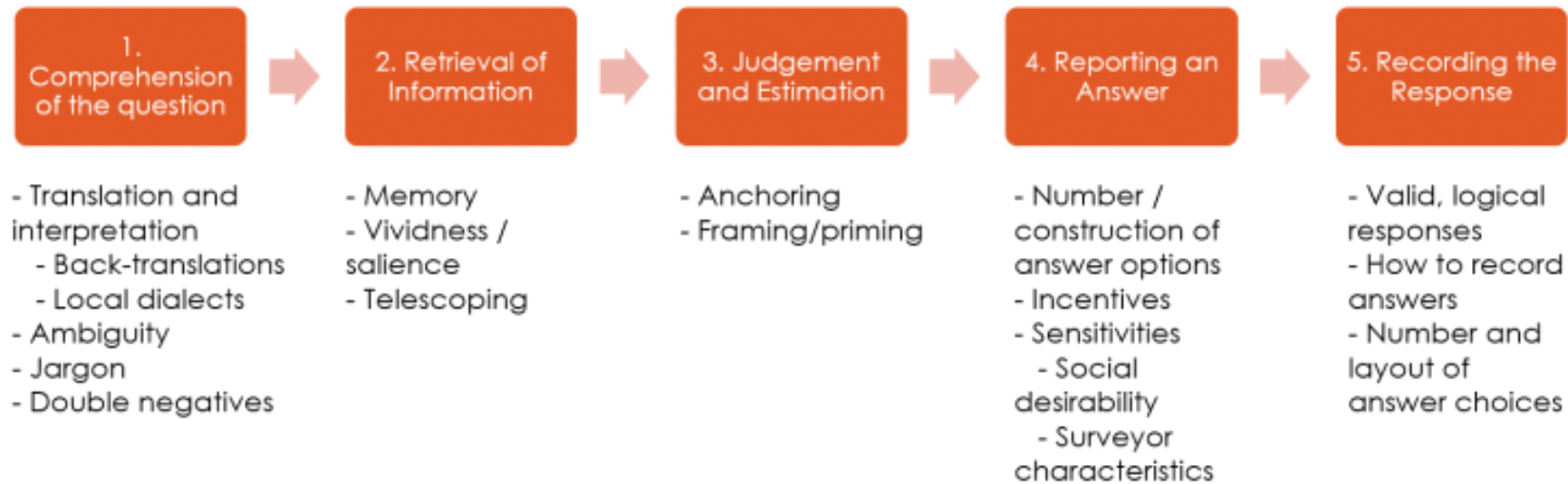




Step 4: Response

1.1 Total monthly income, before taxes





Sources of measurement error

Review your questions for issues

- ▶ See: types of measurement errors
 - ▶ <https://www.povertyactionlab.org/resource/introduction-measurement-and-indicators>

Issue	Example	Tip
<p>Vagueness can result in respondents interpreting questions in different ways.</p>	<p>"How many times did you consume rice this month?"</p> <p>Respondents may have different interpretations of "rice," "times," "consume," and "this month." For example, does "this month" refer to the past 30 days or the month in which the respondent is being interviewed (e.g., the month of June)?</p>	<p>Look at each word in your question carefully, brainstorm alternate meanings and define any ambiguous concept. This is particularly important with abstract concepts, e.g., empowerment, risk aversion, or trust.</p>
<p>Questions that involve negatives can be confusing and lead to misinterpretations</p>	<p>"Many people regularly do not eat at least one meal per week. For how many weeks in the last year was this not true for your family?"</p>	<p>Avoid using negatives wherever possible</p>
<p>Double-barrel questions: When a question has multiple parts, it may not be clear which part the question respondents are answering</p>	<p>"Should the government provide free education because school is too expensive in our community?"</p>	<p>Avoid convoluted sentence structures, and break complex questions into their constituent parts</p>

Question and response types

39

- ▶ Open vs. closed ended
 - ▶ Open-ended question: where respondents provide a response in their own words
 - ▶ Text, numerical
 - ▶ Closed-ended question: where they are asked to choose from a list of answer choices
 - ▶ Can be select-one or select-multiple responses
 - ▶ Make sure responses are:
 - ▶ Clear and simple
 - ▶ Mutually exclusive
 - ▶ Completely exhaustive (include “other (specify)” wherever you are in doubt)
 - ▶ Allocated code numbers consistently across questions
- ▶ Can use more open ended at the pilot stage to see what responses are most common and create closed ended responses accordingly

Response scales

40

- ▶ One type of common closed-ended question uses response scales
 - *Likert scales* (strongly agree, agree, disagree, strongly disagree)
 - *Ranked responses*: Ask respondents to *rank a set of options* through numbering.
 - These help us understand levels of importance and help respondents to differentiate/prioritize,
 - But they also assume that respondents feel different about different options and that they know enough about each item to make a comparative judgment.
 - Ranking takes time and mental effort; in general, you should *limit your list to fewer than 7 options*.

Plan sequencing and skip patterns

41

- ▶ Wherever possible, questions (or questions within a section) should flow as follows:
 - *From general to specific*
 - *From less sensitive to more sensitive.* Early questions should be easy and pleasant to answer and should build rapport between the respondent and the researcher
 - *From factual to attitudinal* in order to build up rapport on objective questions before moving to questions that reveal something about the respondent's beliefs or opinions
- ▶ Ensure that you use skip patterns, particularly when surveys are long. **Skip patterns** direct the flow of questions so that respondents answer only those questions that are relevant to them.
 - ▶ Mistakes with skip patterns can lead to two possible types of errors:
 - *Errors of commission:* Individuals get questions that are not relevant to them, and therefore enter incorrect/misleading information.
 - *Errors of omission:* Individuals do not get the chance to answer questions that were relevant to them.
 - ▶ Very important to both conceptually think through relevance and skip patterns and test them when programming the questionnaire

Pre-test, program, re-test, and revise

42

- ▶ Once you have developed your draft questionnaire, test it out!
 - ▶ **Cognitive interviewing** is the administration of draft survey questions while collecting additional verbal information about the survey responses,
 - ▶ Used to evaluate the quality of the response and/or to help determine whether the question is generating the information that its author intends
 - ▶ Can do cognitive interviewing on a small convenience sample of respondents/similar respondents
 - ▶ Revise confusing points
- ▶ Program your (revised) survey
 - ▶ Details on software options upcoming
 - ▶ Test out the program and make sure it follows what you planned
 - ▶ Have a small convenience sample test the questionnaire
 - ▶ Check the data comes out as you intended
- ▶ Iterate as needed until the questionnaire performs well

Questionnaire length

43

- ▶ Total length of survey administration should be short!
 - ▶ Web survey: 10 minute median, maximum of 20 minutes
 - ▶ Phone survey: 20 minute median, maximum of 30 minutes
- ▶ Must always be prepared to ask, "Is this question really needed?"
 - ▶ Preparing the report template can also help identify questions that won't be used and should be cut
- ▶ Test for length early and often!

Data
collection
digital tools

Why digital tools?

45

- ▶ Surveys can be collected on paper or typed into spreadsheets
 - ▶ Prone to error
 - ▶ Costly and time consuming
- ▶ Best to use digital data collection tools to improve accuracy and efficiency
 - ▶ End result is an analyzable dataset
 - ▶ Can be filled either by enumerators (staff) during a call or directly by respondents
- ▶ Large variety of digital tools with different pros and cons
 - ▶ Will talk through a few examples, but can consider other tools for tracer survey as well, so long as they have similar features

Google forms

46

- ▶ Pros:
 - ▶ Free and unlimited submissions
 - ▶ Intuitive interface for programming
- ▶ Cons:
 - ▶ Generally pretty basic
 - ▶ Skip patterns (conditional questions) somewhat limited
 - ▶ Need a paid plan to get support
 - ▶ Publishes data to google sheets, so will need to download and label data in statistical software

The screenshot displays a Google Form titled "Feedback Form" in a purple-themed interface. At the top, it indicates "Section 1 of 5". The form title "Feedback Form" is centered, with a "Form description" field below it. The first question is "Enter your Name", followed by a "Short answer text" input field. The second question is "Enter your Email Address", also followed by a "Short answer text" input field. The third question is "Which Session did you Attend ?", which is a radio button question with three options: "Session 1 : Getting started with Google Forms by Aryan Irani", "Session 2 : Getting started with Google Sheets by Ben Collins", and "Session 3 : Cloud Computing by Jake". At the bottom, there is a navigation bar with the text "After section 1" and "Continue to next section" with a dropdown arrow.

Survey Monkey

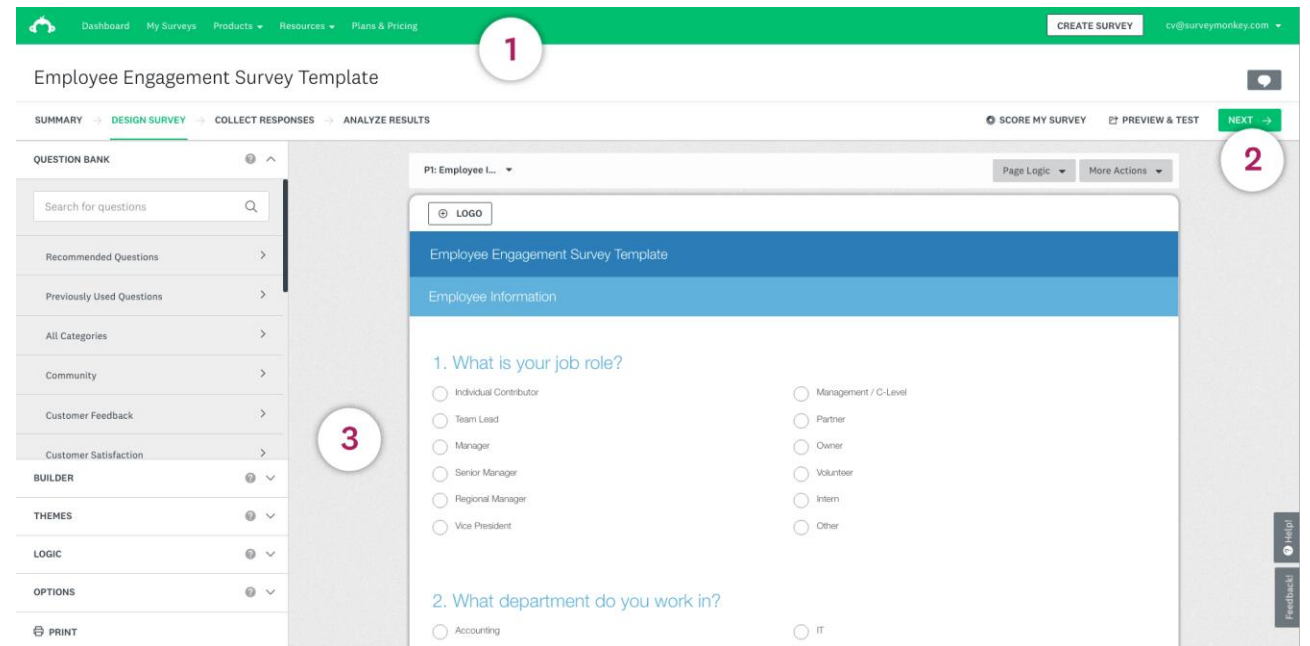
47

▶ Pros:

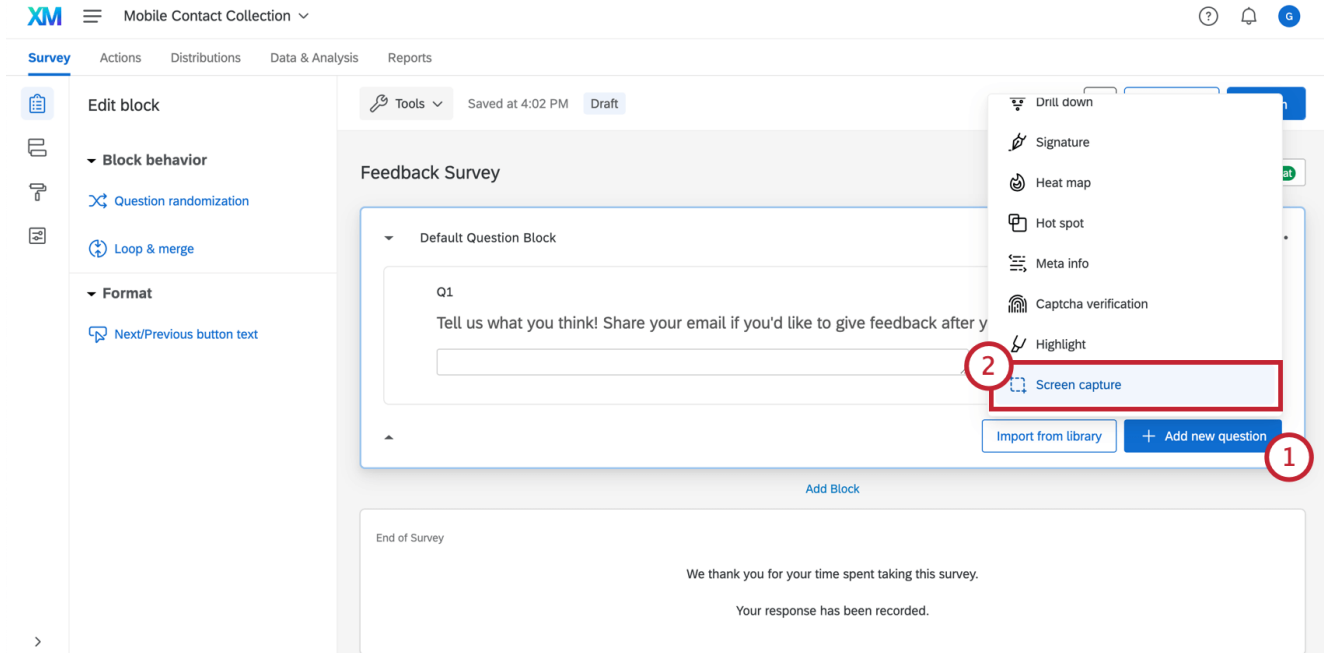
- ▶ More advanced question types, layouts, and routing
- ▶ Provides feedback on length, design
- ▶ Can do response tracking easily
- ▶ Can export to a variety of formats

▶ Cons:

- ▶ Not free, although not super expensive
- ▶ Provides technical support



- ▶ Pros:
 - ▶ More advanced features for survey programming
 - ▶ For example, looping (repeating a set of questions, for example asking the same questions about each job someone has had)
 - ▶ More advanced tools for checking and analyzing data in the platform
- ▶ Cons:
 - ▶ More expensive than SurveyMonkey
 - ▶ May not need many of the advanced features



- ▶ Pros:
 - ▶ Excel-based programming
 - ▶ Additional features
 - ▶ HTML formatting, complex constraints
 - ▶ Good customer support
 - ▶ Commonly used for RCTs
- ▶ Cons:
 - ▶ Complexity
 - ▶ Paid, although may be cheaper than Qualtrics

The next field is a multiple choice field, which uses the *select_one* field type:

type	name	label
select_one yesno	consent	Would you like to continue?

The answer choices to this multiple choice question are listed on the *choices* sheet:

list_name	value	label
yesno	1	Yes
yesno	0	No

For each option in the "yesno" list, there is a *label* (which will appear to users) and a *value* (which is the associated value that will appear in the data). This is how this multiple choice question will appear to users:

Collect > Sample SurveyCTO Form

Would you like to continue?

Yes

No

Questions?

50