Designing Quality Survey Questions

WMEN Meeting

June 16, 2023

EMR 6500: Survey Research

How to design surveys

How to develop and ask survey questions

How to sample in surveys

How to assess survey reliability and validity

How to conduct self-administered internet and mail surveys

How to analyze survey data

How to report on surveys

Checklist for Quality Survey Design

Checklist Item	Yes	No	N/A	Notes
Question type will result in the type of data needed given evaluation questions				
Evaluator is prepared to analyze results from this type of question				
Question asks for information respondent is likely to know, or respondent will have access to information needed				
Question focuses on respondent and respondent experiences (and not what respondent thinks about others)				
Evaluator can anticipate how a respondent might answer the question				
Question relates closely and clearly to one (or more) evaluation question(s)				
Question is as specific and narrowly focused as possible				
Question asks about only one concept (i.e., is not double-barreled)				
Question is concise, using as few words as possible to convey meaning				
Question phrasing is consistent with language used in other places in the survey				
Question is written in semantically valid, plain, straightforward language				
Question is syntactically correct, employing correct grammar and usage				
Question wording reflects the language and culture of respondents				
Question is specific enough for respondents to know exactly what information is being requested				
Question requires as little inference or mental math as possible				

Checklist Item	Yes	No	N/A	Notes
Question is phrased using active voice				
Question is phrased as neutrally as possible				
Question itself includes any necessary instructions for answering				
Explanations are included for any terms that might require clarification				
For an open-ended question, question instructions note how long or detailed answers should be				
Reference periods correspond to the information needed (e.g., shorter time periods for more common occurrences)				
Reference periods are clear and specific (e.g., *since last Tuesday* as opposed to *in the last week*)				
If question is sensitive (e.g., a demographic question), question instructions include information about why the question is being asked and/or how responses will be used/useful				
Question instructions make it clear how to answer (e.g., check all that apply, select the top three)				
Response options are exhaustive; that is, they fully cover the range of expected responses				
Response options are mutually exclusive; that is, they do not overlap				
Response options are labeled (either numerically, pictorially, verbally, or a combination) where appropriate				
Response options are aligned with wording of question stem (i.e., direct labeling or item-specific response options are used)				
Bipolar response options are balanced with the same number of positive and negative options				
5-7 options are used for scaled responses (unless desired level of precision requires fewer or more options)				
Don't know, neutral, N/A or similar responses options are used only as needed and have a clear purpose				

Checklist Item	Yes	No	N/A	Notes
A midpoint is used when a neutral option is needed or when forcing respondents to choose a side may result in measurement error				
The order of scaled response options is consistent (e.g., left to right, least to greatest) with other questions				
Numeric scales are ordered least to greatest				

Ice Breaker

What is your pet peeve when it comes to responding to surveys?

How does that pet peeve influence your response(s) to the survey?

Design Thinking

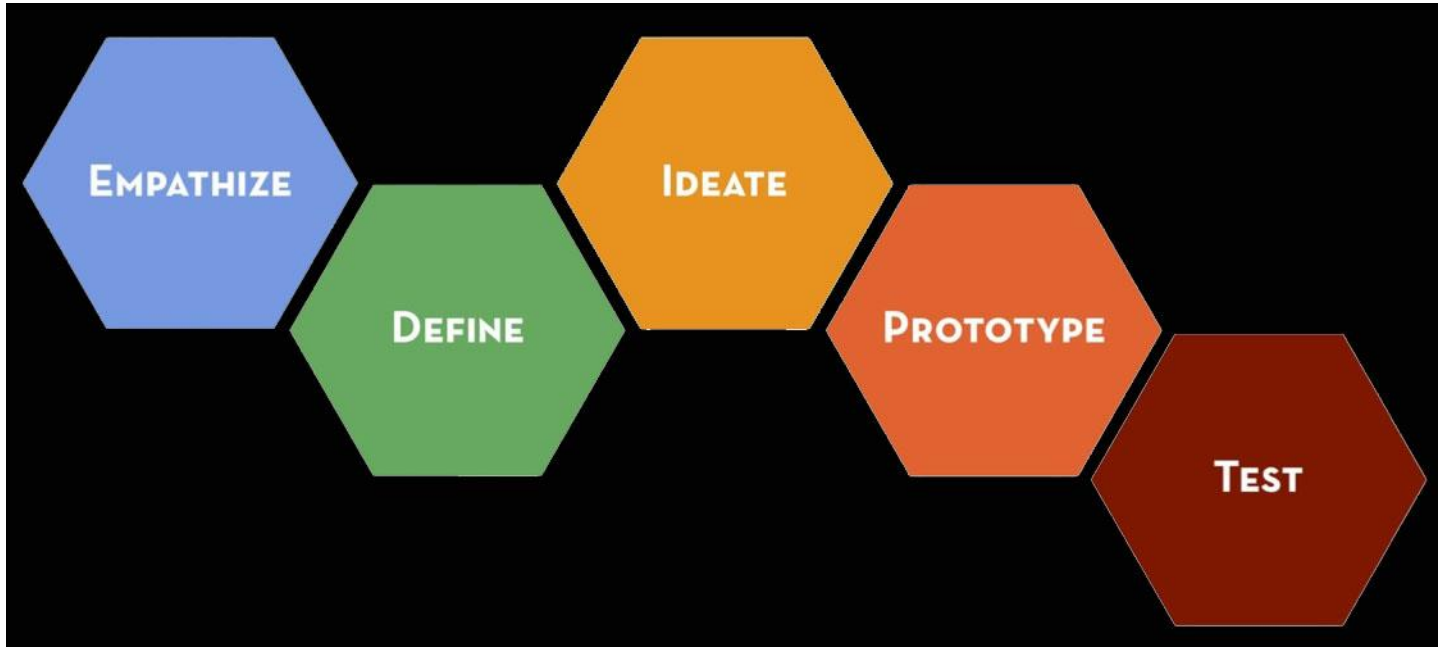
Qualtrics CustomerXM™

Design thinking is a set of phases during which designers attempt to deeply understand end users and their experiences

It is a human-centered approach in which the designer puts ego and self-interest aside to create a product that best meets users' needs

Survey questions that are poorly understood, misinterpreted, or skipped by respondents pose serious problems for the evaluator

Design Thinking



Phase 1: Empathize

Empathy is the centerpiece of a human-centered design process

This is the work you do to understand people, within the context of your design challenge

It is your effort to understand the way they do things and why, their physical and emotional needs, how they think about the world, and what is meaningful to them

Phase 2: Define

This phase is all about bringing clarity and focus to the design space

It is your chance, and responsibility, as a design thinker to define the challenge you are taking on, based on what you have learned about your user and about the context

To take full advantage of this phase, it is helpful to have fully articulated the purpose of the survey

Phase 3: Ideate

To ideate is to generate ideas

This is where we generate and start drafting questions, taking into account what we know about respondents as well as our survey purpose and information needs

Because we have been grounded by the empathize and define phases, we will be able to generate a more focused set of questions which will limit the pruning of ideas and questions in later stages of the process

Phase 4: Prototype

A prototype is created and taken to a set of potential users for their feedback and input

During this phase, we will be ensuring that our survey measures what we want it to measure, i.e., that our questions have face validity

We also begin to ensure that the question stems and response options are appropriately aligned

Phase 5: Test

In survey research, this is often called pretesting

Such testing solicits even more user input to inform potential revision or redesign before a survey is finalized

Common survey pretesting practices include piloting and cognitive interviewing

Purposeful survey design process

Planning the survey

- Determining the purpose of the survey
- Understanding survey respondents (empathize phase of design)
- Understanding what surveys can measure (define phase of design)

Developing the questions

- Sourcing questions (ideate phase of design)
- Crafting question stems, response options, and special consideration questions (prototype phase of design)

Finalizing the survey

- Pretesting (test phase of design)
- Preparing for administration, analysis, and use

Planning the survey

Determining the purpose of the survey

Understanding survey respondents (empathize phase of design)

Understanding what surveys can measure (define phase of design)

Determining the purpose of the survey

First, articulate the evaluation questions that the data to be collected will help answer

Normally, there will be just a few (1-3) evaluation questions, though sub-questions may be used to additional detail

It is important to distinguish these evaluation questions from the questions that are asked in the survey (or in a focus group or key informant interview)

Determining the purpose of the survey

A construct is something that cannot be directly observed (e.g., intelligence, health, prejudice, intent to change, awareness of a social problem, etc.)

An indicator is something that can be observed and also represents a particular construct (e.g., for health, indicators may be heart rate, blood pressure, cholesterol levels, etc.)

Generally speaking, evaluation questions address constructs and survey questions address indicators

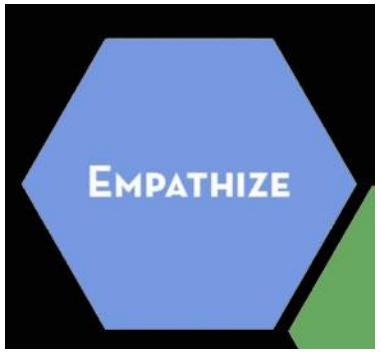
Determining the purpose of the survey

How will the gathered data inform the answers to the evaluation questions?

Who will use the data captured by the survey?

How do we expect the gathered data to be used, including any decisions that rest on the results of the survey?

In what form will the data be needed to maximize its usefulness?

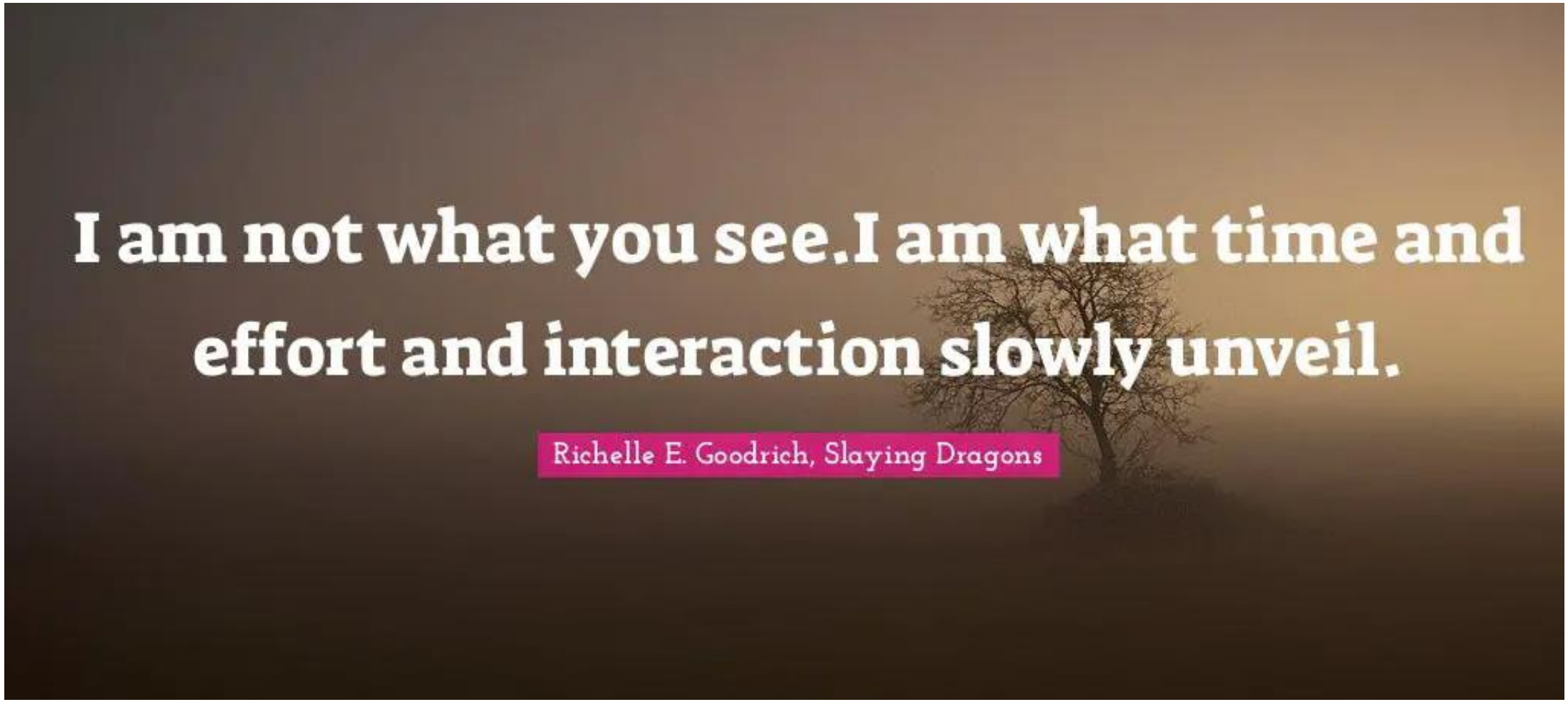


Planning the survey

Determining the purpose of the survey

Understanding survey respondents (empathize phase of design)

Understanding what surveys can measure (define phase of design)


I am not what you see.I am what time and effort and interaction slowly unveil.
Richelle E. Goodrich, Slaying Dragons

Evaluator-centered vs. respondent-centered

An evaluator-centered approach to writing survey questions asks "What do I need to know from respondents?"

A respondent-centered approach to writing survey questions asks

- "What might the respondent want to tell us?"
- "What questions might they be able and willing to answer that would help answer our evaluation question(s)?"

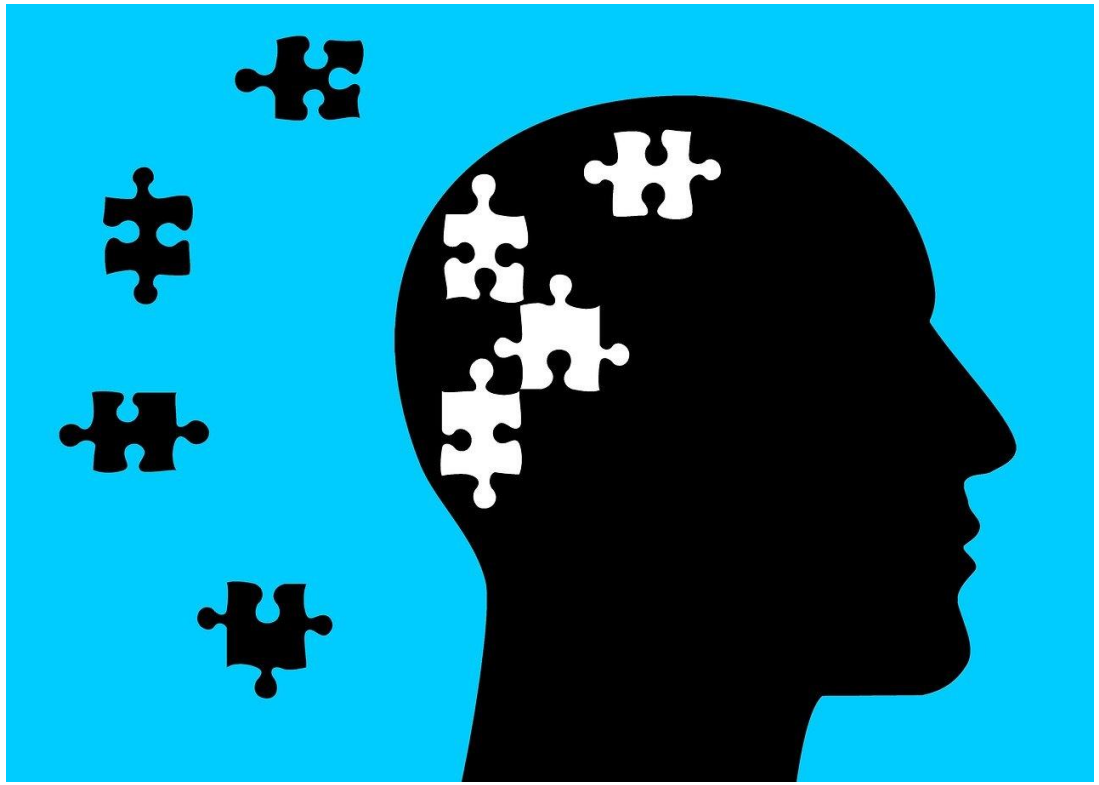
Cognitive tasks required of survey respondents

Comprehension

Retrieval

Judgment

Response

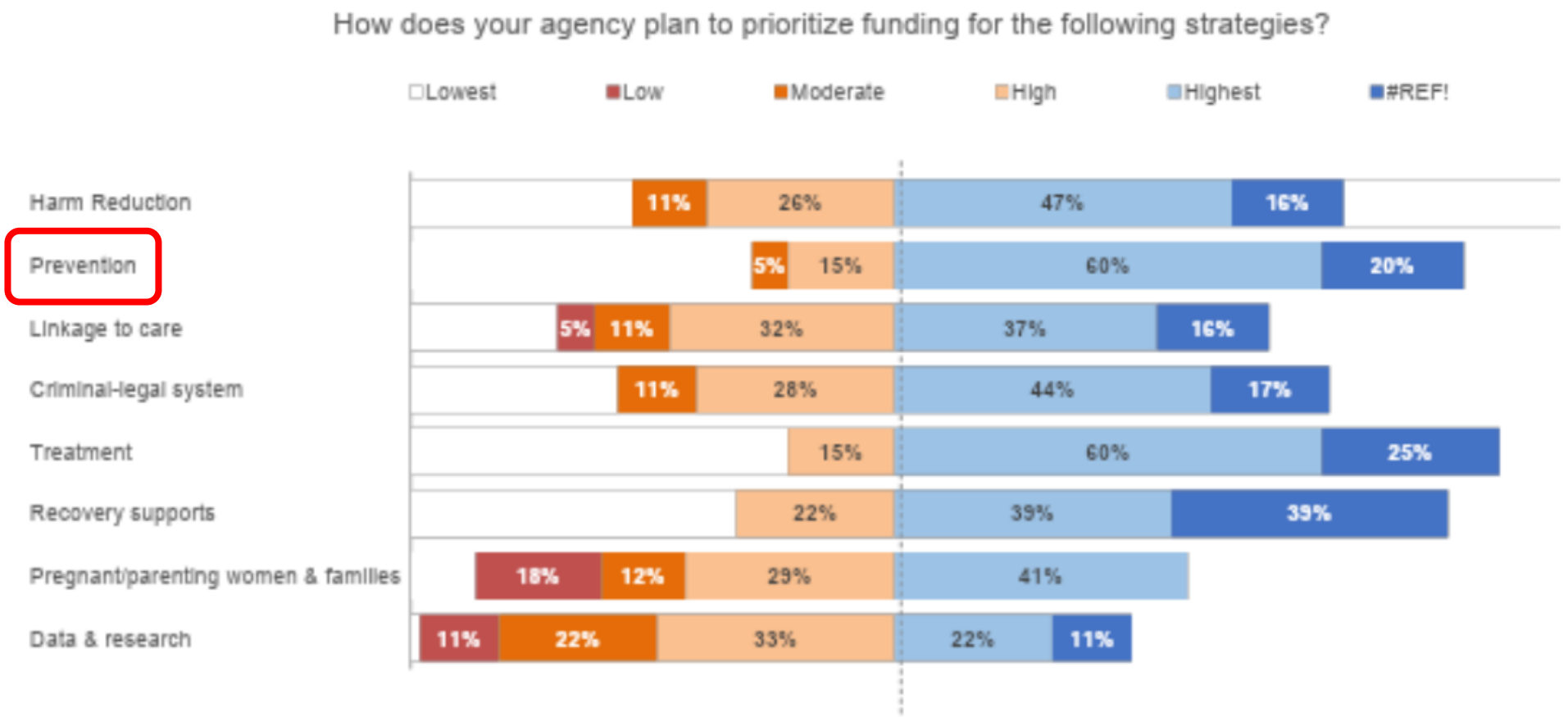


Comprehension

Problems with comprehension arise when respondents

- Do not notice, do not read, or misinterpret instructions
- Encounter unfamiliar vocabulary in a question stem or response option
- Interpret words or phrases differently than the way the evaluator intended
- A question is worded in an overly complex or detailed way

Comprehension



Retrieval

The interplay between survey questions and respondents' memories affects retrieval, including

- The way a question is worded vs. the way in which memories are encoded (e.g., trash pickup vs. garbage collection)
- The degree to which the question provides retrieval cues (e.g., think about trips to the grocery store that you most often visit)
- The passage of time since the event in question and the survey

Judgment

There are three types of judgment

- Judgment for factual questions

For the factual questions, a respondent must determine how accurate their retrieval process is

- Judgment about dates and durations

- Judgment about frequencies

For dates and frequencies, relative mundane events are difficult to remember and distinguish from each other (e.g., does this event count as a visit?)

Response

Responding to a question involves groups of processes around

- “mapping” the answer to available response options (e.g., multiple choice questions) and
- editing the answer to meet certain criteria (e.g., essay question)

Other variables to consider

Respondent willingness

Respondent ability

Bias in language

Context and culture

Respondent willingness

Factors related to willingness include

Question type: Some respondents like open-ended questions because they appreciate the evaluator's interest in their opinion, while others may be reluctant to respond to open-ended questions

Question topic: Questions of a sensitive nature will also impact respondent willingness and motivation to answer; question sensitivity is highly culture and context dependent

Social desirability index: a respondent's tendency to edit their answers to make themselves "look good"

Respondent ability

Two key factors related to respondent ability to accurately answer survey questions include

- Comprehensibility of the question: Questions that are easier to understand are better able to elicit accurate answers; complex questions can result in satisficing answers which involves providing minimally acceptable answers (this could include “don’t know” answers)

- Respondent memory: memories change each time they are recalled and that recall is not accurately depicted by simplistic metaphors such as retrieving a file from a file drawer; **remembering is actually an act of reconstruction**

Bias in language

Accounting for bias requires that we

Understand that there are important differences in the worldviews of key stakeholders

Challenge stereotypes and patterns of marginalization and subordination

Accurately reflect how individuals view their own group memberships

Promote full participation by conducting evaluation activities in the participant’s preferred language

Context and culture

Gone are the assumptions that applied research and evaluation can be entirely objective; the design of data collection tools must attend to context and culture

Context includes political, environmental, organizational, and cultural contexts

Sensitivity and responsiveness to respondents must include an awareness of the difference between the evaluator's culture and the respondent's culture

Context and culture

Cultural competence is about the ability to work across cultures, in a manner that is both respectful and responsive to all people to learn more fully from those involved in the evaluation

Cultural groupings are ascribed differential status and power, with some holding privilege that they may not be aware of and some being relegated to the status of “other”

Cultural privilege can foster disparate treatment in resource distribution and access

Culturally Responsive Evaluation (CRE)

CRE is a holistic framework for centering evaluation in culture

It recognizes that culturally defined values and beliefs lie at the heart of any evaluation effort

Evaluation must be designed and carried out in a way that is responsive to these values and beliefs, many of which may be context specific

Figure 1. Culturally Responsive Evaluation Framework

Context and culture

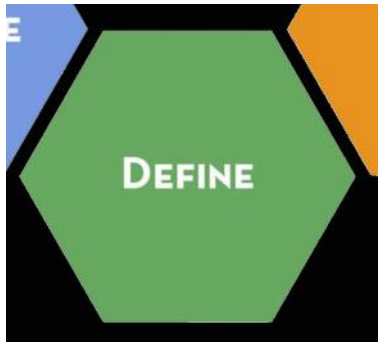
Preparing for Cultural Responsiveness

Considerable preliminary fieldwork is critical to carefully contextualize marginal communities

Document the process of attempts to be culturally responsive at multiple phases of the evaluation

Not only use a culturally responsive data collection instrument, but also administrate the instrument in a culturally responsive way

Acknowledge and address the potential tension between conventional methods of quantitative methodology and the desire to use CRE



Planning the survey

Determining the purpose of the survey

Understanding survey respondents (empathize phase of design)

Understanding what surveys can measure (define phase of design)

Understanding what surveys can measure

Surveys typically capture self-reported transformative information, including

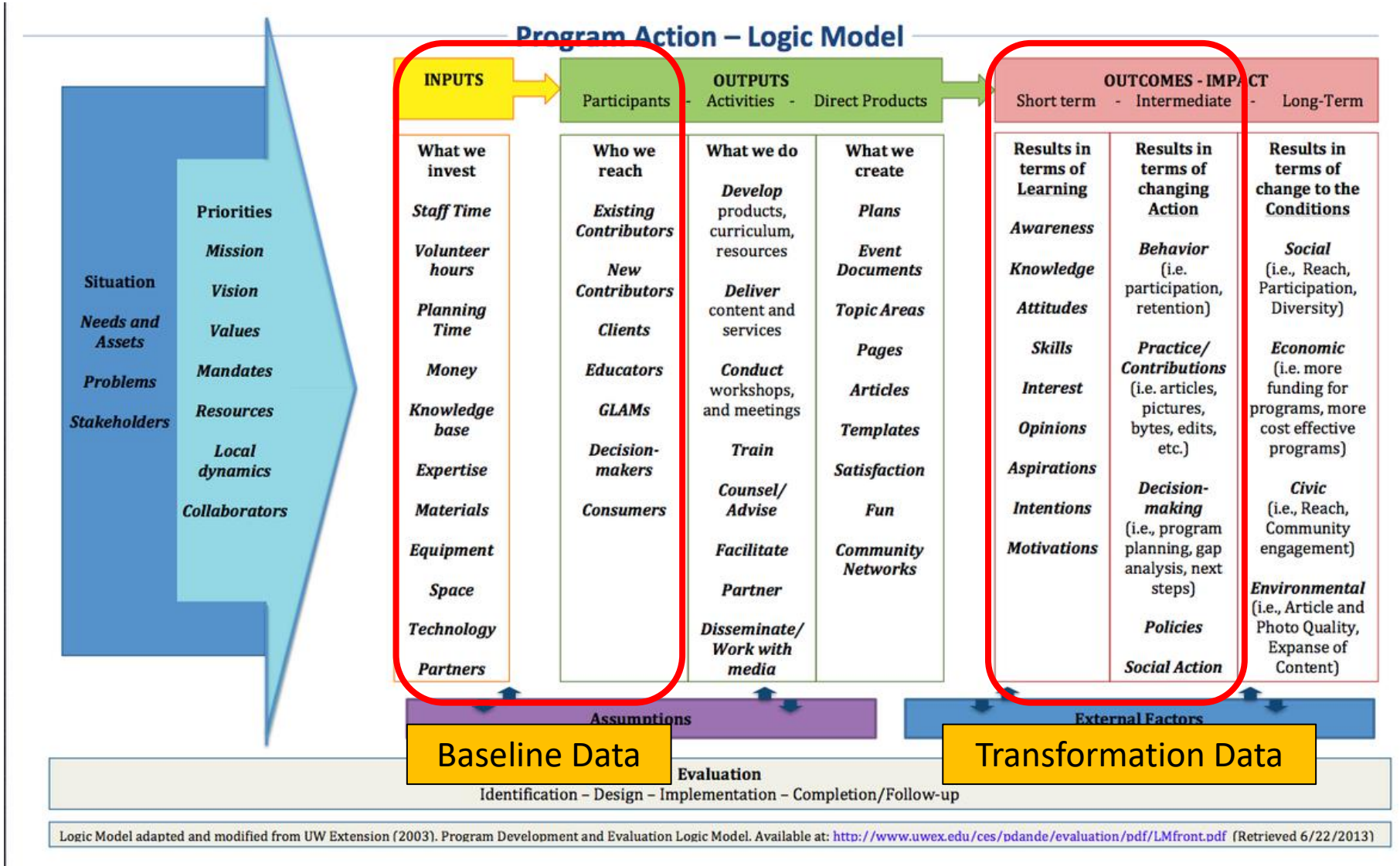
Attributes: Inherent characteristics or qualities such as age, race, ethnicity, nationality, education, employment, religion, gender identity, and sexual orientation

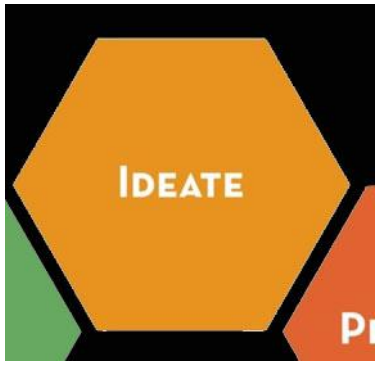
Behaviors: What people actually do, including how often or how many times respondents regularly engage or have engaged in certain behaviors of interest

Abilities: What people are able to do, including skills or knowledge of a particular subject

Thoughts: Perspectives on a topic, including attitudes, beliefs, feelings, awareness, opinions, or preferences about a particular matter

Understanding what surveys can measure





Developing the questions

Sourcing questions (ideate phase of design)

Crafting question stems, response options, and special consideration questions (prototype phase of design)

Brainstorming

Evaluators should not develop survey questions in a vacuum. Instead, engage the help of experts or those in the target population of respondents

Brainstorming sessions with others provides inspiration for potential questions and serves as a sounding board for existing questions

Having gone through the design thinking phases of empathizing and defining, a brainstorming session for sourcing questions tends to be more focused

Brainstorming

In this phase, designers use the question stem “How might we…” For ideating (sourcing) survey questions, we might ask:

- How might we ask respondents about their health habits?
- How might we understand their experiences at the event?
- How might we figure out if they intend to change their practice?
- How might we ask questions of respondents in ways that limit bias?

Engaging potential respondents or informants

Engaging potential respondents in interviews or focus groups provide input well worth the additional time and resources required

Evaluators typically invite small groups (6-10 people) to engage in a moderated discussion of the survey topic

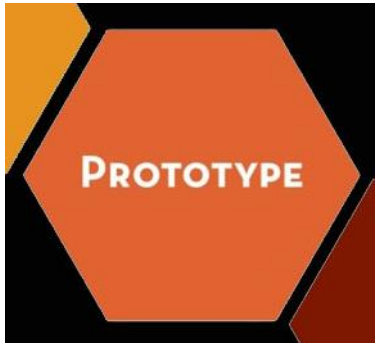
Individual interviews are effective in soliciting input from informants with subject-matter expertise

These focus groups and interviews can also be used in the pretesting phase once a survey draft has been created

Using or adapting existing measures

Evaluators can and should explore existing surveys or measures before creating their own. This is a good approach when

- The survey topic has been well studied by others
- Existing tools are accessible
- Existing tools have undergone validation and reliability tests



Developing the questions

Sourcing questions (ideate phase of design)

Crafting question stems, response options, and special consideration questions (prototype phase of design)

Crafting question stems

- Reference periods
- Problematic text features
- Ambiguous wording
- Double-barreled questions
- Leading and loaded questions
- Open-ended questions

Reference periods

Reference periods are the specific time frames evaluators ask about in individual survey questions (e.g., in the last month, or in the last year)

Both shorter and longer reference periods possess inherent advantages

- Because memories fade with time, respondents may have an easier time searching withing shorter time periods
- However, if respondents rarely experience the feeling or behavior of interest, the reference period must be long enough to result in useful responses

Problematic text features

- Acronyms
- Low-frequency terms, i.e., words less often encountered, e.g., use *angry* instead of *irate*
- Vague quantification terms, i.e., words used to communicate uncertain or approximate amounts, e.g., exact numbers instead of frequently, many, few, etc.
- Left-embedded syntactic structures, i.e., the respondent must get through several phrases, adjectives, etc. before they get to the critical part of the question

Problematic text features

- Ambiguous syntactic structures, i.e., the structure of the sentence or phrase is open to more than one interpretation
- Dense noun phrases, i.e., phrases where many adjectives, adverbs, or articles are attached to nouns
- Quantitative mental calculations, e.g., about how many hours per week…
- Hypothetical questions, i.e., about assumed situations as opposed to current facts, i.e., what if questions

Problematic text features

- Numerous logical operators, i.e., connecting words such as *and, but, or*, and *if*—too many in one question taxes the working memory
- Nominalizations, i.e., verbs or adjectives that have been changed into nouns, e.g., instead of the phrase: *the expansion of*, use *to expand*
- Passive construction, i.e., when the object of the action in a sentence becomes the subject, e.g., *The store was expanded by the developers* vs. *The developers expanded the store*
- Bridging inferences, i.e., when respondents must make inferences about an introductory sentence in order to answer the question

Ambiguous wording

Ambiguous wording makes questions difficult for respondents to answer accurately

Use the QUAID (Question Understanding Aid) tool to get feedback on question wording and sentence structure, including

- Unfamiliar technical terms
- Vague or imprecise noun phrases
- Complex syntax
- Working memory overload

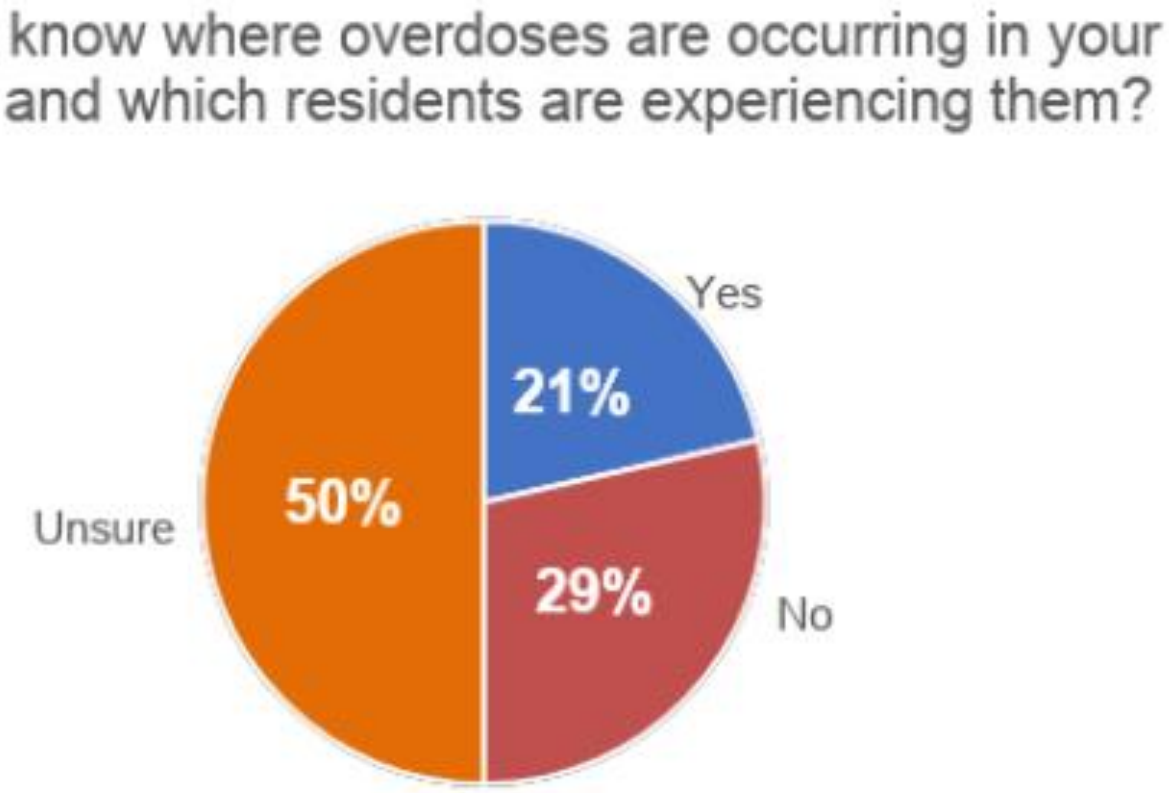
Double-barreled questions

Questions must ask respondents only one thing at a time

Example of a double-barreled question: *Rate how knowledgeable and helpful we were in assisting you*.

Example of a triple-barreled question: *Has the district been responsive to your customer service needs with timely, informative, and accurate information?*

Double-barreled questions



Leading and loaded questions

Leading questions can lead respondents to provide certain answers, e.g., *"How much did you enjoy the movie"* uses language that implies the respondents did indeed enjoy the movie.

Loaded questions can influence question responses, especially when political views are being studied, e.g., consider these pairs

- *Pro-choice vs. pro-abortion*
- *Pro-life vs. anti-abortion*
- *Taxes vs. revenue*

Leading and loaded questions

However, both leading and loaded questions can be used strategically

Some advocate for the deliberate use of these questions to increase the likelihood that respondents will report on certain sensitive behaviors (e.g., *“It’s OK if you do this…”* or *“Many people do this…”)*

The use of these types of questions requires a deep understanding of the survey topic and of the desired respondents to know that such a strategy would be more helpful than harmful

Open-ended questions

Open-ended questions are used when there is

- Too wide a range of potential responses to construct a reasonable list of response options
- Uncertainty about an appropriate set of response options
- A need to capture rich, detailed, and nuanced understanding of respondents' thoughts, feelings, attitudes, or experiences
- A need to hear the respondents answers in their own voice, using word choices and terminology that is comfortable for them

Open-ended questions

The best open-ended questions are simple and highly specific, letting respondents know what exactly we hope to learn from them

- Tell them why the question is being asked
- Use introductory language to put fears to rest about individual writing styles
- Break longer or complex questions into separate parts
- Use at the end of a survey to invite more sharing of information, e.g., *“Is there anything else that you would like to share about…”*

Crafting response options

- How many response options to include
- The odds or evens debate
- Unipolar and bipolar response scales
- Using a “Don’t Know” option
- Frequency scales
- Likert and Likert-type items
- Matching response options to question stems

How many response options to include

There is no right or wrong answer to the question of exactly how many response options to offer for any given survey question

In theory, the more response options to choose from, the more a respondent will be able to distinguish her attitude toward one thing versus another

The downside of offering many options is that the difference between various points can get messy and the meaning of any specific point may be less precise

The odds or evens debate

Including a mid-point option (which results in an odd number of responses) is a matter of ongoing debate

Many evaluators feel that few if any respondents are genuinely neutral about a topic and therefore they should not be allowed to choose a mid-point (neutral) option

Some reasons that a respondent would choose a neutral option include satisficing, lack of certainty on a topic, social desirability and other biases, or the avoidance of the mental effort required to select another response option

Whether a mid-point is used or not, there should be a balance of positive and negative response options

Unipolar and bipolar response scales

Unipolar scales contain levels of a specific construct (e.g., *understanding*) ranging from the absence of the construct (e.g. *no understanding*) to a high degree of the construct (e.g., *great understanding*)

Bipolar rating scales measure a construct from one polar opposite to the other (e.g., *dissatisfaction vs. satisfaction*)

Some constructs lend themselves more readily to either a unipolar or bipolar scale. Unipolar constructs have a natural zero and theoretically cannot go below zero (e.g., *effectiveness ranges from not effective to very effective*)

Unipolar and bipolar response scales

FIGURE 5.4 ● Example of a Survey Question With a Unipolar Rating Scale

Please rate your level of understanding of quantum physics:

- O I have no understanding of quantum physics
- O I have a little bit of understanding of quantum physics
- O I have a fair amount of understanding of quantum physics
- O I have a great deal of understanding of quantum physics

FIGURE 5.5 ● Example of a Survey Question With a Bipolar Rating Scale

Please rate your overall satisfaction with our services:

- O Very dissatisfied
- O Somewhat dissatisfied
- O Neither satisfied nor dissatisfied
- O Somewhat satisfied
- O Very satisfied

Unipolar and bipolar response scales

FIGURE 5.6 ● Real-World Questions: *Poorly Constructed Bipolar Scale*

I and/or my organization have used the following {organization name} services and resources within the last 5 years:

	Frequently	Somewhat frequently	Not sure	Less frequently	Infrequently
Keynote speeches	○	○	○	○	○
Tools/workbooks	○	○	○	○	○
Infographic posters	○	○	○	○	○

Bipolar

A suggested alternative: Revise question stem and use unipolar response options with an added "not sure" option (at the end)

How often would you say you or your organization used {organization name} services and resources in the last 5 years?

	Very frequently	Somewhat frequently	Infrequently	Not at all	Not sure
Keynote speeches	○	○	○	○	○
Tools/workbooks	○	○	○	○	○
Infographic posters	○	○	○	○	○

Unipolar

Using a “Don’t Know” option

It is important to anticipate that respondents may not know the answer to every question, regardless of how well we have empathized with and developed understanding of our desired respondents

At times *“don’t know”* provides meaningful data, like the respondent truly does not know the answer, or they may find the question to be too sensitive, or the question may be written poorly

A filter question could be used to determine if the respondent knows enough or is interested enough about the topic to have formed an opinion

Using a “Don’t Know” option

FIGURE 5.8 ● Using a Filter Question Instead of a “Don’t Know” Option

1. Do you feel as if you have enough information on the war in Syria to have formed an opinion on U.S. involvement in it?

   ○ YES

   ○ NO

If you answered YES, please proceed to question 2. If you answered NO, please skip to question 3.

Frequency scales

Ordered sets of response options called frequency scales can be used to provide options for *“how often”* or *“how many times”* a respondent engaged in a behavior, felt an emotion, or had a particular experience

It is important to understand how the frequencies offered (i.e*., low frequencies or high frequencies*) might affect how the respondent answers the question

For example, the question on the following slide asks about frequency of sadness. If low frequency options are offered, the respondent may interpret the question to be asking about major sadness, and vice versa.

Frequency scales

FIGURE 5.9 ● Interpreting a Question Stem in the Context of a Frequency Scale

Respondents will likely interpret "sadness" differently in these two questions:

Version 1: How often do you feel sad?

- O About once per year
- O A few times per year
- O At least once per month
- O At least once per week

Version 2: How often do you feel sad?

- O About once per month
- O A few times per month
- O At least once per week
- O Several times per week
- O At least once per day

Likert and Likert-type items

A Likert scale is a cluster of survey items related to each other that probes a particular construct of interest with very specific set of consistent response options

A Likert-like scale looks like a Likert scale but has one or more variances associated with it

It is probably best to not use the term “Likert” and refer to these types of questions as closed-ended questions with a rating scale or with an agree-disagree scale

Likert scale

FIGURE 5.10 ● Example Format of a Likert Scale

Please rate the degree to which you agree or disagree with each statement.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
High-quality public education is one of the most important issues facing Americans today.	1	2	3	4	5
Every American deserves access to a free and appropriate public education.	1	2	3	4	5
America currently has the best public school system in the world.	1	2	3	4	5

Likert-type scale

FIGURE 5.12 ● A Likert-Type Item

Please rate your level of satisfaction with:

	Dissatisfied	Somewhat Satisfied	Satisfied	Very Satisfied
The customer service you received.	1	2	3	4

Crafting response options (cont.)

- Matching response options to question stems
- Offering enough response options
- Labeling response options
- Ordering response options
- Decomposition
- Retrieval cues

Matching response options to question stems

Response options must match the question being asked

The best way to avoid mismatching is to label response options with the same language that matches what the question is asking

- Scales with matching language are called direct labeling scales or item-specific scales

Matching response options to question stems

FIGURE 5.17 ● Examples of Matching Question Stem and Rating Scale Language (Item-Specific Rating Scales)

How helpful was our staff?

- O Not at all helpful
- O Somewhat helpful
- O Very helpful

Please rate the performance of our staff.

- O Terrible
- O Poor
- O Fair
- O Good
- O Very good
- O Excellent

How useful is the product?

- O Not at all useful
- O Somewhat useful
- O Very useful
- O Extremely useful

Please rate your level of satisfaction with the product.

- O Very dissatisfied
- O Somewhat dissatisfied
- O Neither satisfied nor dissatisfied
- O Somewhat satisfied
- O Very satisfied

Offering enough response options

Ensure that response options reflect the entire range of possible response, i.e., they must be exhaustive

However, trying to cram in all of the possible options may make for too many choices, resulting in a heavy cognitive load for respondents

For some questions, ensuring sufficient response options without overburdening the respondent can be alleviated by using a catch all response such as “other”

Labeling response options

Another ongoing debate is whether to use verbal labels or just numbers on a rating scale

Labeling just the end points of the scale with numbers in the middle eases the cognitive load for respondents, but makes it more likely that respondents with an extreme response style (ERS) will choose only the endpoints

Studies have shown greater benefit with applying verbal labels to all scale points

Also, using the most specific and precise language is preferable to using vague quantifiers (e.g., many, most, sometimes, etc.)

Ordering response options

Findings of studies on response option order are inconsistent

There is some evidence of a primacy effect where the first items seen are best remembered and most influential

Order effects may be most evident in respondents with less cognitive sophistication

Also, most respondents tend to lean toward the positive end of the scale (positivity bias) and avoid selecting scale endpoints (response contraction)

Ordering response options

Robinson and Leonard offer this advice for ordering response options

- It is best to randomize the order of response options and then randomly assign respondents to different versions of the question
- If a numeric scale is used, it should read from least to greatest, either left to right or top to bottom
- Rating scales should generally start with least desirable to most desirable to mitigate the primacy effect

Decomposition

Decomposition is breaking down complex questions into smaller chunks by asking multiple questions about subcategories of the event of interest, providing retrieval cues along a dimension such as time, place, person, etc.

The intention of using retrieval cues is to help the respondent remember instances of an event that may otherwise be forgotten

Decomposition is probably best used when the reference period is shorter

Decomposition

FIGURE 5.24 ● Decomposing a Bipolar Rating Scale Question

Undecomposed question:

Overall, how satisfied are you with today's purchase?

○	○	○	○	○
Very satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Very dissatisfied

Decomposed question:

Overall, are you satisfied, dissatisfied, or neutral about today's purchase?

○	○	○
Satisfied	Neutral	Dissatisfied

Overall are you slightly, somewhat, or very satisfied/dissatisfied?

○	○	○
Slightly	Somewhat	Very

Retrieval cues

Retrieval cues appear in survey questions as definitions or examples of what is being asked. Advice for using cues include

- It is better to give examples that cover the full category range
- The best examples are those that are likely to be relevant and likely to be overlooked
- Words may be better cues than pictures

Retrieval cues

FIGURE 5.25 ● Survey Question Featuring a Definition and Examples as Retrieval Cues

How many times in the last 12 months did you participate in an individual sport competition? Individual sports are those played alone, without teammates. Examples include bowling, boxing, cycling, golf, skiing, swimming, running, and wrestling.

Special consideration questions

- Filter questions
- Sensitive questions
- Demographic questions

Filter questions

Filter questions (AKA screening questions) can be used to

- Screen respondents for eligibility to complete the survey
- Direct respondents to different questions or sets of questions, depending on their answers to the filter question
- Solicit additional information about particular responses
- Avoid the use of a “don’t know” response option

Sensitive questions

Sensitive questions are those that feel intrusive or invasive to the respondent. The categories of questions may be considered sensitive by the respondent.

Financial matters

Mental or technical ability

Self-perceived shortcomings

Social status

Sexual identity or sexual behavior

Alcohol consumption or illegal drug use

Emotional or psychological disturbance

Sensitive questions

Sensitive questions may elicit a social desirability bias in the respondent

To mitigate the effect of social desirability bias, use closed-ended questions

Use a wider range of response options that might otherwise be constructed for non-sensitive questions

Use forgiving language to encourage respondents to be forthcoming in their responses to sensitive questions (see next slide)

Sensitive questions

FIGURE 6.5 ● Real-World Questions: *Using Forgiving Language in a Sensitive Question*

In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. Which of the following statements best describes you?

- ○ I did not vote (in the election this November).
- ○ I thought about voting this time, but didn't.
- ○ I usually vote, but didn't this time.
- ○ I am sure I voted.

Source: ANES (2016).

Demographic questions

These questions help the evaluator describe the attributes of a population such as age, gender, sexual orientation, etc.

These questions are also used to identify sub-groups for use in the analysis of a survey

It is important to limit the number and type of demographic questions included in a survey; they should be included only if they serve a clear purpose

These questions may compromise the anonymity of a survey, especially with smaller populations

Demographic questions

Respondents may feel alienated or even offended if they don't find themselves represented in the response options

If a *"Prefer not to answer"* option is not offered, respondents may not be truthful or may give up on the survey

Other considerations for demographic questions are how to ask them, where to place them in a survey, and how to create a positive rapport to encourage them to willingly give answers

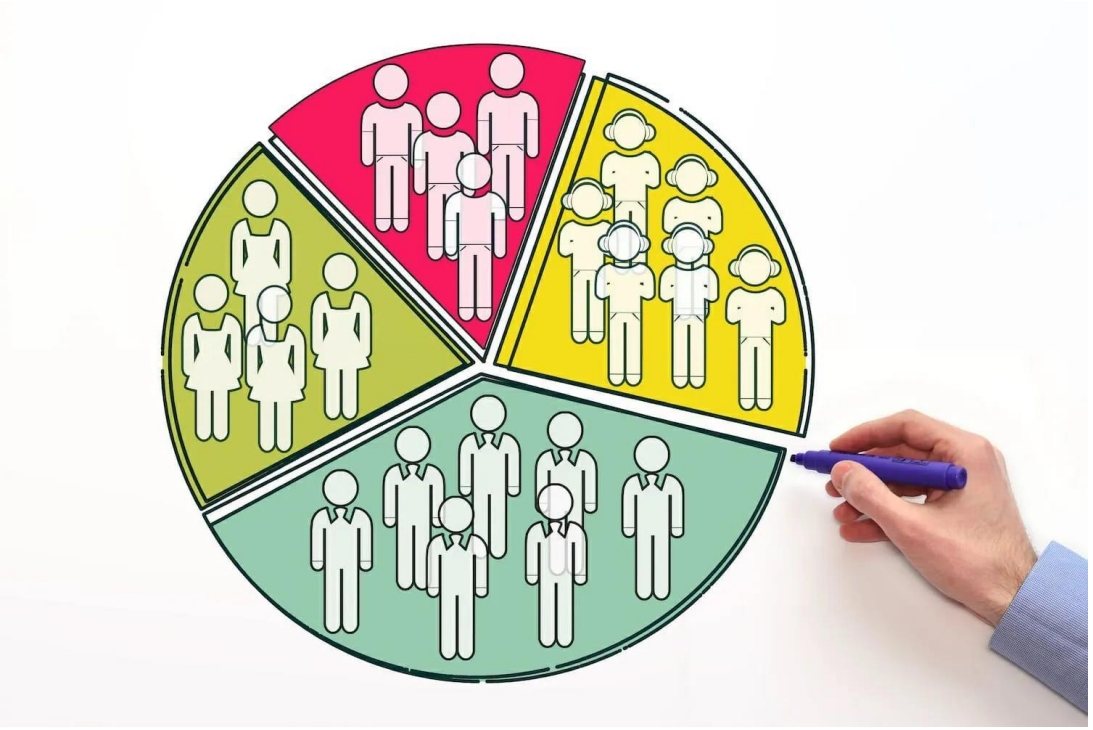
Demographic questions

Too many demographic questions can contribute to respondent fatigue, especially if placed at the beginning of the survey

Still, since some demographic questions are easy to answer, these questions could be at the beginning of a survey to get respondents warmed up for upcoming questions that require more attention and effort

A main objective is to get respondents engaged with the survey as quickly as possible and starting with the most relevant and interesting questions is an effective way to do so

Specific demographic questions



Asking about age

Asking about race/ethnicity

Asking about gender identification and/or sexual preference

Asking about income

Asking about age

The format of the age question makes relatively little difference to item non-response rates, though providing a series of age ranges as response options may be helpful

Regarding accuracy, it is advised to ask respondents to give their birthdate or year they were born

Asking about race/ethnicity

Though race and ethnicity are often used interchangeably in surveys, they are not the same

Ethnicity refers to a person's ancestry, heritage, or culture, whereas race is a socially constructed category with no evidence of a biological or genetic basis

These terms carry with them both history (e.g., colonization, oppression, marginalization) and implicit or explicit assumptions (e.g., stereotypes) about the people identified by these terms

The authors stopped short of making recommendations and strongly suggest doing our own research for the most up-to-date resources (including consulting with potential respondents)

Asking about gender identification and/or sexual preference

There is no “right way” of asking these questions

Gender identity and sexual preference are not the same and sexual preferences are not limited to hetero- or homosexuality

If these types of questions are used, they must serve a specific purpose and that purpose must be used to design the specific questions

Include the reason for asking the questions along with an explanation of how resulting data will be used

These questions should be at the end of the demographics section

Asking about gender identification and/or sexual preference

While the terms gender and sex are easily confused, they refer to who we are and our biological makeup, respectively

The terms man and woman are gender terms, while female and male describe sex

It is increasingly common to offer the *“prefer not to answer”* option. A *“prefer to describe”* may also be an acceptable option.

The word “other” can carry a negative connotation and make people feel excluded and should not be used

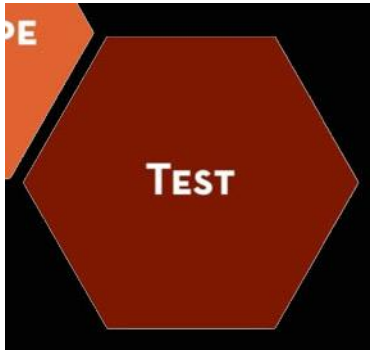
Asking about income

Income data may be important when studying the effects of social stratification, inequality, and poverty

At times, all that is needed is an understanding of the sufficiency of income—whether it is a source of worry or stress

Socioeconomic status (SES) encompasses not only income but also educational attainment, financial security, and subjective perceptions of social status and social class

Asking education level may be a proxy (though not perfect) for income



Finalizing the survey

Pretesting (test phase of design)

Preparing for administration, analysis, and use

Advantages of pretesting

Gather others' input to gain new perspective

Develop greater understanding and empathy with desired respondents

Ensure questions are clear, comprehensible, and relevant to respondents

Confirm that response options are sufficient and appropriate and that respondents can use them consistently (for reliability)

Explore whether questions have face validity, i.e., it appears to measure what it is intended to measure

Anticipate how the resulting data might be analyzed

Expert review

Experts can provide a fresh look at draft survey tools and are often able to ask questions about language choices or otherwise challenge assumptions we might not realize that we are making

The greatest risk in soliciting only expert review is that experts are not at all guaranteed to think like desired respondents

Cognitive interviewing

Cognitive interviewing (CI) is intended to help evaluators understand more deeply how respondents think about the survey questions

Participants are often chosen to represent desired respondents

The process is often iterative; multiple rounds of CI may be necessary

The two most common CI strategies are

- Asking participants to think aloud either as or after they review questions
- Asking probing questions about the participants' experience in understanding and responding to a question

Focus groups

Focus groups can be used to design survey questions and test survey questions

Differences between focus groups and cognitive interviewing include

- Focus groups (FG) are used to explore topics more generally rather than review participant's thought processes about questions
- FG protocols are likely to be more flexible than CI protocols
- Because FGs are conducted socially, interaction is allowed and encouraged which may result in rich information surfacing

Pilot testing

In a true pilot test, a small but strategic sample of desired respondents is drawn and asked to complete a draft of the survey

Pilot tests can

- Determine needs for survey instructions
- Identify questions that are problematic
- Determine preferences regarding question order or identify other order issues (e.g., skip logic problems)
- Identify changes needed in the survey process itself, e.g., more time is needed to complete the survey, invitations are not being received correctly, etc.

Finalizing the survey

Pretesting (test phase of design)

Preparing for administration, analysis, and use

Survey instructions

Keep instructions as simple and brief as possible while also giving the respondents the information they need

There should be a short page of contextual information that precedes the survey and mirrors the language in the survey invitation

This preamble should include the purpose of the survey, the deadline for responding, any info needed to answer the questions, and contact info in case the respondent has questions or technical challenges

Headings at the top of each section of questions should provide context specific to those questions, e.g., *"These questions in this section ask you to share information about…"*

Order effects

In general, surveys should be organized with logical flow that is likely to make sense to the respondent

Order effects occur when answers to later questions on a survey are influenced by earlier questions so that possibility should be considered

Place at least some of the most interesting or relevant (to the respondents) questions early in the survey

Begin with less sensitive or threatening questions, e.g., it is usually best to put sensitive demographic questions toward the end

Begin with questions that are relatively easy to answer

Group questions that are related into meaningful categories or topics, though there situations where the order of questions should be randomized

Formatting surveys

Respondents draw meaning from nonverbal as well as verbal cues, e.g., making inferences about the words in the response options, the shape of the scale, and even the spacing of the options

The following are five interpretive heuristics respondents may use

- Middle means typical (or neutral)
- Left and top mean first
- Near means related (e.g., keep stem and options close together)
- Up means good
- Like (in appearance) means close (in meaning)

Formatting surveys

Visual text features such as bold text, italics, underline, boxes, indentation, and color should be used with caution because they may distract respondents more than assist them

For open-ended questions, a larger answer box may result in an increase in the length and quality of responses, especially for less motivated respondents

For questions that have responses with a midpoint, it may be helpful to separate a non-substantive response option (e.g., no opinion) from the rest of the options to avoid a shifting of the midpoint

Formatting surveys

FIGURE 8.2 ● Visual Midpoint Interferes With Conceptual Midpoint

How would you rate the quality of the materials you received?

- ○ Excellent
- ○ Good
- ○ Fair
- ○ Poor
- ○ No opinion

Separating the nonsubstantive option from the substantive options:

How would you rate the quality of the materials you received?

- ○ Excellent
- ○ Good
- ○ Fair
- ○ Poor

- ○ No opinion

Formatting surveys

One study showed that using a drop-down feature for a question about age resulted in significant response error. The authors concluded that drop-down response menus should be avoided when possible.

It is increasingly important to consider what surveys will look like on a smartphone or other mobile device, e.g., Qualtrics uses page breaks to present questions one at a time on smartphones

Formatting surveys for visual disabilities

Use a 12-point font size or larger

Select colors with a high level of contrast (e.g., black type on white paper)

Avoid too much formatting; use bold for emphasis rather than italics, which are more difficult to read

Provide a narrative description for any images, tables, or charts used and avoid images of text, e.g., avoid using scanned documents

Blind Foundation in New Zealand (2018)

Reviewing for reading comprehension

Two tests of readability in Microsoft Word

- **Flesch Reading Ease Test**
  - Produces a score up to 100, with higher scores indicating easier reading
- **Flesch-Kincaid Grade Level**
  - Tests approximate grade level (US). Readable.io is an online tool that serves a similar purpose.
  - Aim for at least **a grade level or two below** where most respondents are likely to be. Conventional wisdom calls for no higher than 9th grade reading level for the general public (average reading level in the US for adults is 7th to 8th grade)
  - Surveys on more technical subjects (e.g., health-related issues) may benefit from **a lower grade reading (5th grade or lower)**

Contacting potential respondents

Principles of persuasion

Compelling invitations

Sufficient and engaging reminders

Points of contact

Informed consent/assent

Survey incentives

Qualtrics
CustomerXM™

Principles of persuasion

Reciprocation: People try to repay in kind what another person provides

Commitment and consistency: Once someone has committed, there will be pressure to be consistent with that commitment

Social proof: We view a behavior as correct in a given situation to the degree that we see others performing it

Liking: People usually say yes to the requests of someone they like

Authority: Most people have learned that obedience to proper authority is right and disobedience is wrong

Scarcity: Opportunities seem more valuable when their availability is limited

Influence: The Six Principles of Persuasion, Robert Cialdini

Compelling invitations

A comprehensive invitation typically includes

- Who is conducting the survey (draws on the principle of authority)
- The intended use of the survey
- The deadline for response
- Why the respondent should care about responding (using incentives draws on the principle of reciprocation)
- Promises of anonymity or confidentiality (remember, these are different)
- Any detail need to ensure informed consent/assent
- Any information they might need to help them respond
- Researcher contact information

A personal tone draws on the principle of liking

An emotional appeal for response draws on the principle of commitment and consistency

Sufficient and engaging reminders

- A brief overview of the purpose of the survey and why responses are needed
- Instructions for how to respond and link to the survey
- Incentives
- Deadline for response
- Contact information of the evaluator

It may be helpful to thank all those who have responded (draws on the principle of social proof)

It may also be helpful to plan for extensions by extending the deadline for a few days (draws on the principle of scarcity)

Points of contact

Most survey efforts at the very least include

- An invitation
- One or more reminders
- A thank-you or acknowledgement

For an online survey, 4-5 points of contact are likely needed: an initial invitation, 3-4 reminders, and an acknowledgement

Informed consent/assent

Regardless of what is required more formally, the following basic principles of formal consent can be followed:

- Provide a complete picture of the survey's purpose
- Let respondents know whether responses will be confidential or anonymous
- Inform respondents how the data will be stored, used, and shared
- Disclose who will have access to the information collected

Survey incentives

The benefits of responding ought to outweigh the costs experienced

To increase the benefits

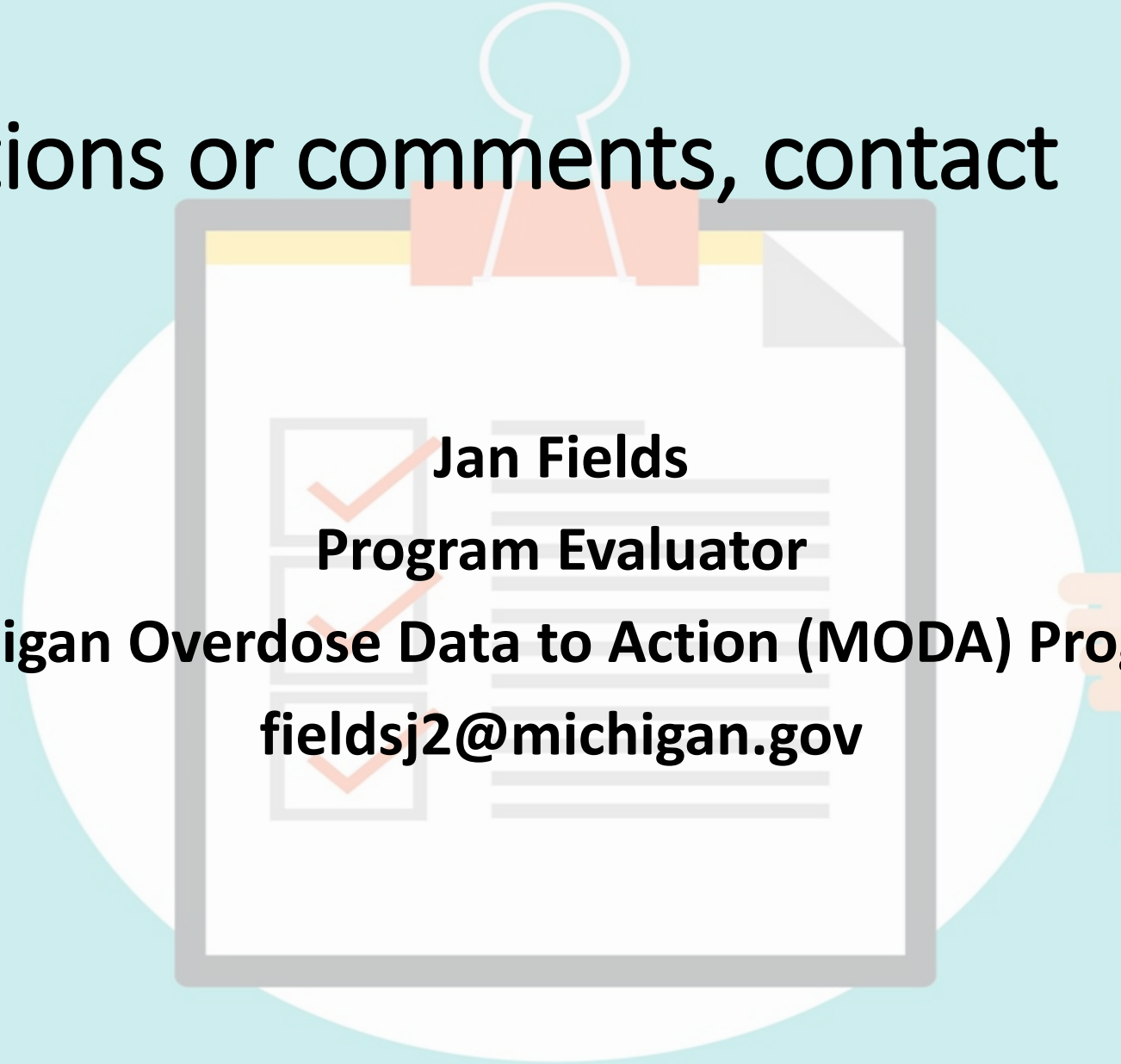
- Specify how the results will be useful
- Pique the respondent's interest by asking interesting questions
- Stress that opportunities to respond are limited
- Use cash or material incentives to encourage, but not require, reciprocity

Survey incentives

To reduce the cost

- Reduce the length and complexity of the survey
- Use visual design principles to make questionnaires easier to complete
- Avoid subordinating, alienating, or excluding language
- Avoid collection of sensitive or threatening information unless absolutely necessary to answer evaluation questions

For questions or comments, contact



Jan Fields

Program Evaluator

Michigan Overdose Data to Action (MODA) Program

fieldsj2@michigan.gov