Table of Contents

1. Designing from multiple perspectives
2. Preparation
   • Defining and operationalizing a research question
   • Sampling considerations
   • Attrition & non-response
   • Ethical considerations
3. Drafting your survey
   • Question wording
   • Question types & response options
   • Data quality checks
   • Question order
4. Reviewing your draft
There are three “people” or perspectives whose needs you must consider:

- **The participant**
- **You as question asker**
- **You as results interpreter**

The needs of these three people will often be in opposition and must be balanced.
Defining & Operationalizing a Research Question

Not defining research questions first can lead to...

1. Missing a key question
2. Measuring a key question in a way that makes analysis or interpretation difficult
3. A ‘kitchen-sink’ approach to question inclusion
EXAMPLE: Research Questions
What happens when we skip this step?

Instead of defining RQs, maybe you describe general survey goals:

- To understand whether this event was worth doing
- To understand how we could improve it next time

The outcome may be a survey like this, which seems fine at a glance:

- How satisfied were you with the event overall? ...with the speaker? ...with the location? ...with the networking opportunities? ...with the Q&A format?
- How many people did you network with at this event? Were your questions answered? If you hadn’t attended, how much of this information could you have learned another way?

- How likely are you to recommend this event to other people? How likely are you to attend another event like this in the future? Why or why not?
- Which topics were most useful to you? Which were least useful? Do you have any suggestions for improving the event? Who would you like to see as a speaker in the future?

But:

- Have we missed a key question?
- How are we going to interpret them to understand whether the event was worth doing?
- Time-consuming
Defining & Operationalizing a Research Question

1. Define your Research Question
   • State clearly, in a short phrase, the main thing or things (3 max) that you want to know
   • And possible answers: Jot down possible answers to your question (informal hypotheses and alternative hypotheses)

2. Operationalize your Research Question
   • Then (and only then) attempt to answer the question above with one or more concrete measures in your survey
   • Which specific outcomes on the survey indicate an answer of X vs. Y?
   • If you need to compare groups in order to answer your question (e.g. people under vs. over 35), make sure you include a question to let you determine which group they’re in
EXAMPLE: Research Questions
Following the Define & Operationalize Process

Let’s transform these objectives into Research Questions:

• To understand whether this event was worth doing
• To understand how we could improve it next time

➢ Interrogate your objectives: What does “worth doing” mean? What does improvement mean? Use this to turn them into more specific RQs.

Possible Research Questions:

• Was this event considered useful by most attendees? By non-attendees?
• Was this event more or less useful than other events we’ve run (and could run again)?
• Did this event meaningfully change attendees’ understanding of a topic?
• Are there ways we could make it more useful to attendees next time?
• Are there ways we could increase attendance (especially for any underrepresented segments of the group) next time?
Survey Length

• Most first survey drafts are too long
• Have someone unfamiliar with the content pilot test for length

Overall recommendations:
• Keep it as short as possible, especially if participants are unpaid
• If you only need one question, only ask one question

By type:
• Program evaluation: < 5 min. (1-2 min. if feasible)
• Unpaid attitude/behavior surveys: < 5 min. (1-2 min. if feasible)
• Paid attitude/behavior surveys: 10-15 min. (25-30 min. at the most)
Sampling Considerations

- Surveys are usually completed by a **sample** of the **population of interest**
- First, carefully define that population to help with survey distribution
  - People subscribed to your email list...
  - All subscribers, including ones who sign up but never open emails? Just more active ones? Just those who have been on the list longer than ___?
  - Screener questions can be used to establish your criteria (never assume everyone will carefully read your invitation letter)
- It is preferable to know:
  - Population size: for calculating response rate (# completes/# population)
  - Basic population characteristics (e.g., demographics, other info relevant to what you’re measuring): for understanding representativeness and examining non-response bias
- Sample sizes
  - Use a **sample size calculator** to calculate how many people you need for a representative survey or a **power analysis** to calculate how many you need to find a significant difference between two groups
Attrition & Non-Response

• Definitions
  • Attrition = partial survey response (dropping out partway through)
  • Non-response = not starting the survey (low response rate = a lot of non-response)

• People often ignore non-response/partial response because they think it’s not a big deal or that there’s nothing they can do about it. Neither is true.
  • Random non-response isn’t that bad, but differential non-response is common and can have major implications for interpreting results
  • Prevention: Design your invitation and consent form to minimize non-response and your survey to minimize attrition. Measure basic population characteristics if possible (e.g., during event signup: gender, age, first event?)
  • Reaction: If your data is already collected...
Attrition & Non-Response

(*cont’d*) Reaction: If your data is already collected...

1. If you have a basic demographic breakdown of your population (e.g., 45% of people who signed up for this event were male, 70% were 18-25, 55% were first-timers)
   • Compare these percentages against your sample: how do the gender, age and first-time breakdowns compare? What does it mean for interpretation if they’re substantially different?

2. If you don’t have specific population numbers but have some other type of evidence (e.g., at an in-person event, you observed that about two-thirds were female and a show of hands suggested about half were attending for the first time)
   • Compare these percentages against your sample. What does it mean for interpretation if your estimates and the sample breakdown are substantially different?
   • Use any available evidence as long as you report the source and any subjectivity

3. If you know nothing about the population:
   • Describe how any plausible differences between your sample and population would affect your overall results and the conclusions you draw from them (e.g., it’s usually safe to assume that people who are less engaged in a group are less likely to participate in a survey about the group)
   • Consider the worst-case scenario: e.g., if every non-respondent had selected “not at all satisfied” with your event, how would your overall results look?
Attrition & Non-Response: Causes

The best solution is to keep attrition and non-response as low as you can.

Some Potential Causes of Non-Response

• Estimated time investment is too long
• Low intrinsic motivation and no incentive to participate
• Lack of certainty about anonymity or confidentiality, esp. if questions are sensitive or participant has negative/controversial opinions to express
• No reminder sent (busy people may forget)
• **Solutions**: keep it short, have informed consent, use reminders, have an incentive

Some Potential Causes of Attrition

• Long surveys, esp. without incentives
• Questions that are hard to understand or answer
• Too many open-ended questions
• Long matrix/grid-style questions
• Formatting or usability issues, esp. on mobile
• **Solutions**: keep it short, review for problematic questions & formatting, be attuned to things that annoy you as a participant in other surveys
Attrition & Non-Response: Types

Random
- If you ignore response rate, you are implicitly assuming that any non-response was random
- A percentage of people don’t take (or finish) your survey, but they don’t differ from the rest on any meaningful characteristics
- Lowers response rate but isn’t especially problematic

Differential
- A percentage of people don’t take (or finish) your survey, and this is more true of some demographics than others
- This is much more problematic because it introduces bias (also called non-response bias). Your results are no longer true of your population, only of the ones who complete the survey.
- Common low-response groups include:
  - Less interested or motivated people
  - People from less privileged groups

Unknown
- If you don’t have even a rough idea of your population demographics, you won’t know which of the above you have, or how much
- Differential attrition is common, so this is a substantial problem
Ethical Considerations

Familiarize yourself with the core principles of ethical research with human participants:

1. **Respect for Persons**: Respect participants’ autonomy. A key component is seeking free, informed, and ongoing consent.

2. **Concern for Welfare**: Protect welfare of participants in view of any foreseeable risks associated with the research, including physical, psychological, and reputational harm.

3. **Justice**: Treat people fairly and equitably, so that no segment of the population is unduly burdened by the harms of research or denied the benefits of the knowledge generated from it.

To learn more:

1. For an overview, see [Faunalytics’ Research Ethics and Data Handling Policy](#).
2. Most non-profit researchers have no legal obligation to do so, but anyone conducting more than a single survey with no sensitive questions should take the [TCPS-2 CORE](#), a 2-hour online course.
Ethical Considerations: Consent Checklist

- A brief description of the project purpose and procedure
- Expected risks, inconveniences, and benefits (if any)
- A description of any incentives for participating
- Information about how to withdraw from the study
- Information about the confidentiality, use, and retention of data
  - Definitions: *Anonymous* data means no one knows or could reconstruct who provided it. If you collect email addresses, IP addresses, or enough other info that you can sometimes tell who said what, you can treat the data as *confidential* but it is not anonymous, even if you *anonymize* it by deleting that information later.
  - Be clear with participants about who will have access to the data, whether it will be summarized or raw, and confidential or anonymous.
- Researcher contact information
- A statement of consent
You are invited to complete this survey because you watched Faunalytics’ survey design presentation. Your participation is voluntary, but your responses will help us evaluate our impact and improve our programs. The survey is estimated to take about 5 minutes to complete.

Please note that your responses are confidential but will be seen by members of the Faunalytics research team. Please ensure that you do not provide any identifying information if you would prefer to remain anonymous.

You can indicate that you consent to participate by clicking next and completing the survey.
Designing Your Survey
Question Wording: Guidelines

- Keep your questions short, simple, and easy to read
  - A good question can be understood with one quick readthrough by someone without specialist knowledge (unless you are sure that every participant has it)
  - Do you need to provide extra information? If so, make it a separate sentence so they get the most important part in that short, simple, easy-to-read format. This bullet point is an example.

- Don’t target the average member of your population, target someone around the 10th percentile of:
  - Focus
  - Reading ability
  - Motivation to do a good job
  - Conscientiousness

- Always have someone else review and test your survey
## Question Wording: Issues Checklist

<table>
<thead>
<tr>
<th>Check (good/bad indicated with +/-)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>❑ Is it double-barrelled? That is, are there two questions hidden in one? (-)</td>
<td>• Participants will have a hard time answering and you’ll have a hard time interpreting the answers</td>
</tr>
</tbody>
</table>
| ❑ Is it leading? i.e., does it list one option when there are really several, or imply that one option is better than another? (-) | • You may affect how participants answer  
• Results will be open to legitimate criticism |
| ❑ Does it have a negation? Or worse, multiple negations? (-) | • Small negation words like ‘not’ or ‘don’t’ are easily overlooked by participants  
• Multiple negations are confusing  
• These count even when part of other words (e.g., I dislike people who avoid eating meat) |
| ❑ Is it straightforward and specific? e.g., short, declarative, active voice, concrete, no conjunctions and clauses (+) | • Easy to read = more people answer accurately  
• Hard to read = less motivated or more time-pressed participants will answer randomly, skip it, or even drop out of the survey |
| ❑ Would a 13-year-old or non-native speaker be able to answer it? (+) | • Similar to above, but watch out for words that some participants may not know |
| ❑ (If using multiple questions to get at a single concept) Do you have roughly equal numbers of positively- and negative-worded questions? (+) | • Balances out any positivity or negativity bias  
• Try to make negatively-worded questions very clear, and consider emphasizing the negative with bold or caps |
EXAMPLE
Event Question Wording

1. What percentage of sessions at the Faunalytics Capacity-Building Workshop did you find valuable? ______

2. Strongly disagree (1) - Strongly agree (5)
   - The Faunalytics Capacity-Building Workshop had a welcoming environment
   - The Faunalytics Capacity-Building Workshop was an environment where people were open to exploring strategies that are different than those they already use and believe in
   - The Faunalytics Capacity-Building Workshop was an environment where people talked about their advocacy methods with humility
   - The Faunalytics Capacity-Building Workshop taught me about effective approaches I can try

3. Why did you decide not to sign up for a follow-up session?
   - There was no session that was relevant to me
   - The session I wanted to participate in didn’t have any room available
   - I didn’t have time
   - I wasn’t aware of the sessions
   - Other
EXAMPLE

Suggestions to Improve Question Wording

1. Overall, how valuable did you find the sessions at the Faunalytics Capacity-Building Workshop?
   Not at all valuable (1) - Extremely valuable (5) [and N/A if needed]

2. Strongly disagree (1) - Strongly agree (5)
   • I felt welcome at the Faunalytics Capacity-Building Workshop
   • People at the Faunalytics Capacity-Building Workshop were open to exploring new strategies
   • People at the Faunalytics Capacity-Building Workshop were overconfident in their methods
   • The Faunalytics Capacity-Building Workshop taught me about effective approaches I can try

3. Why didn’t you sign up for a follow-up session? Select the biggest reason.
   • There was no session that was relevant to me
   • The session I wanted to participate in didn’t have any room available
   • I didn’t have time to meet [to sign up?]
   • I wasn’t aware of the sessions
   • Other (please specify) _____________________
Question Types & Response Options

Common Types:

1. Likert scale
   (including matrix/grid format)
2. Ranking questions [not recommended]
3. Categorical (one option)
4. Categorical (select all)
5. Open-ended
Likert Scales

- Ordinal: order of options matters, but the attitudinal “distance” between each option is not necessarily the same.
- Very flexible: Can be used to assess attitudes, beliefs, emotions, likelihood of a behavior, etc.
- How many points?
  - 5-7 is generally recommended, as it allows for variance but not confusion/effort.
- Should you label all the points?
  - Easier to understand and report, but can be difficult (esp. with asymmetrical scales) and forces an interpretation. Always label at least the endpoints and (if applicable) the midpoint.
- Symmetrical or asymmetrical?
  - Symmetrical: Very unlikely – very likely
  - Asymmetrical: Not at all likely – extremely likely
  - Asymmetrical gives more room to one end of the scale than the other—harder to interpret, but useful if you expect a lot more responses on (e.g.) the positive end.
- Do you want a midpoint?
  - Lets people be neutral, which can be good or bad. There is no neutral point on an asymmetrical scale.
- Use standard response options if possible
- Avoid long/frequent matrix questions
- Avoid leading instructions by saying (e.g.) “do you agree or disagree with...” rather than just “do you agree with...”
Ranking Questions

- Can be difficult/frustrating for participants
  - Drag-and-drop can be physically challenging or not work in some browsers
  - Number entry is easy to do wrong
  - Forces them to rank options that may not feel meaningfully different to them

- Not great for analysis:
  - The meaning of each person’s rankings depends on how they ranked other things; once averaged across participants, may not mean much
  - Rank is purely relative—no information about how much they prefer one option over another

- Bottom line: use Likert scales instead
  - e.g., very low priority (1) to very high priority (7)
  - Doesn’t force people to rank things they feel are equal
  - Putting the averages in order provides a ranking
  - Tells you more about the distance between options
Categorical Scales

- Order of options is not meaningful
- Many demographics are categorical: gender, race, region, occupation, etc.
- Sometimes there’s a choice between “select one” and “select all”:
  - “Select one” is easier for analysis and interpretation (sometimes much more)
  - “Select all” is more flexible for participants
  - Think carefully about whether more than one option will apply for many people
- Include responses that capture any hypotheses you have and all competing hypotheses
- Generally provide an ‘other (please specify)’ option to catch anything you didn’t think of
- Provide a ‘none of the above’ option if that is a possible answer
- Randomize order of response options for each participant if possible
  - Except other/none of the above—put last
Open-Ended Questions

• Easiest type of question to ask
• Most likely to be skipped or cause survey attrition (dropout)
• Can provide rich detail
• Hardest to analyze if you want to quantify responses
• Use sparingly, and only if both are true:
  • Other question types cannot be used—e.g., if you don’t know enough about the possible answers to create options, and
  • The question is important enough to warrant participants’ time
• Avoid using near the beginning of the survey
If Faunalytics held another Capacity-Building Virtual Retreat, how likely would you be to recommend:

[Not at all likely (0) - Extremely likely (10)]

- Watching live sessions
- Watching recorded sessions
- Submitting questions for Q&A
- Participating in the live text chat during sessions
- Attending one-hour workshops
- Joining speed meetings
- Text chatting with attendees
- Finding new connections with attendees
- Scheduling virtual meetings
- Holding virtual meetings
EXAMPLE: Other Options

**Simplified Likert scale**
If Faunalytics held another Virtual Retreat, how likely would you be to recommend each of the following to another attendee:

[1 - Not at all likely -- 5 - Extremely likely]
- Watching live sessions
- Watching recorded sessions
- [etc.... cut down the number of items if possible]

**Categorical scale (select all)**
If Faunalytics held another Virtual Retreat, **which** of the following would you recommend to another attendee? Select all that apply.
- Watching live sessions
- Watching recorded sessions
- [etc.]
- Another feature of the conference (please specify)
- None of the above

**Categorical scale (select one)**
If Faunalytics held another Virtual Retreat, **which** of the following would you recommend as the best feature?
- Watching live sessions
- Watching recorded sessions
- [etc.]
- Another feature of the conference (please specify)

**Open-ended**
If Faunalytics held another Virtual Retreat, **what** would you recommend as the best features of the conference?
Longitudinal or Repeated Surveys

Definitions

• Longitudinal = same participants at multiple time points (e.g., before and after an event), with the ability to link data (e.g., which before score goes with which after score)
• Repeated = a survey or survey question that is asked the same way each month, year, etc., but without the participants necessarily being the same and without data linking

For either type, comparison requires consistent question wording, response options, and survey administration method.

Common Pitfalls

1. Tweaking once you’ve collected data: Is it worth trading possibly improved accuracy for a definite inability to compare with last year’s responses?
   • If you have to make a change, try to make it a “nested” change so that you can still compare. E.g., “social media” → “Facebook”, “Twitter”, “other social media”. Or if it makes sense to do so, include the old question and the new one.

2. Asking participants to self-report a change when you can calculate the change
   • E.g., ask “Which of these foods do you eat?” both times and calculate the score difference
   • Not “Did seeing this documentary make you eat less of these foods?”
EXAMPLE

Changing a Question on a Repeated Survey

This is a real question from U.S. polling: Do you feel that the country, in general, is headed in the right direction, or do you think things are seriously off on the wrong track?

- Right direction
- Wrong track
- Neither
- Unsure

How could this question be improved?

- Shorter?
- Less idiomatic?
- Use the same wording for positive and negative options?

But should it be changed?

- This question has been used for decades
- Comparability is the main goal
- As a result, it continues to be used today
- At the same time, some polling groups have changed the wording in attempts to improve it, with the result that there are many different variations in use today and it’s hard to compare across them

Do it well* the first time, then leave it alone.

*Note that this says well, not right. Every survey question has flaws. Is your tweak trading one flaw for another?
Data Quality Checks

• Many researchers include questions or other checks to catch people who are either deliberately trying to defraud you or just not paying close attention—both introduce random error into your results.

• These may or may not be necessary for your survey. The following factors all increase the necessity:
  • Offering a participation incentive
  • Recruiting participants from the general population (vs. from an EA group or animal advocacy org, etc.)
  • Online data collection

• Excluding data for failing one data quality check is rarely a good idea, as there are usually extenuating circumstances that can account for one failure.

• We often adopt a 3-strikes rule, but the number of failures is affected by how many checks were included.
## Data Quality Checks: Overview

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Examples</th>
<th>Pros</th>
<th>Cons*</th>
</tr>
</thead>
</table>
| Hidden        | Different records with same IP address  
• IP addresses from outside expected area (e.g., outside U.S.)  
• Very fast completions (e.g., < one third of median time) | • Completely unobtrusive  
• No added burden for participants | • Makes data non-anonymous if collecting IP addresses  
• More time-consuming to analyze than some other options |
| Explicit      | “Choose somewhat disagree for this question.” [on a Likert question] | • Easy to analyze | • These questions may affect participants’ other answers (e.g., due to a feeling that you are trying to “catch them out”) |
| Attention     | In the past month, have you...  
• Used the internet*  
• Ran a mile in under 2 minutes**  
• [among distractor items]  
*mandatory for online surveys  
**impossible; world record = 3:43 | • Subtle | • Select-all questions can be tricky to analyze if participants have the option of skipping the question |
| Consistency   | • How old are you?  
• [many other questions in between]  
• What year were you born? | • Subtle | • Easy to analyze |
| Open-Ended Question Coding | Require at least a sentence in response to an open-ended question. Check whether each response makes sense and reads well. See link below for specifics. | • Subtle  
• Best quality check we are aware of (based on experience & reported success) | • Coding responses can be time-consuming |

We have not listed “could exclude honest responders” as a con each time. It is true of all of them, so no single check should be used in isolation. Learn more about data quality checks at [https://faunalytics.org/data-quality-assurance-plan/](https://faunalytics.org/data-quality-assurance-plan/)
Question Order

• **There is no hard-and-fast answer to the best question order**

• **General guidelines:**
  • Put important questions near the beginning to maximize attention
  • Put objective questions (e.g., demographics) at the end of the survey except as needed for screening—these can influence other Qs but not be influenced
  • Closed- and open-ended questions on the same topic: put open-ended first or responses will be influenced by the closed-ended options
  • Maintain logical flow (e.g., chronological) of questions when possible

• **Counterbalancing or randomizing question order:**
  • Counterbalancing = you create multiple survey versions with different Q order (e.g., ABC, CBA)
  • Randomizing = survey platform randomly assigns each participant to a Q order (including all possible orders (this is the better option if available)
  • Usually for a subset of questions, not the whole survey
  • Consider when you have questions that could influence each other in either direction or any time you have a long section of similar questions—even if there are no clear influences between them (similar to randomizing order of response options when categorical)
Reviewing Your Draft
Reviewing your Draft

Order and prioritization of the three perspectives:

SECOND BUT MOST IMPORTANT
The participant

FIRST
You as question asker

THIRD
You as results interpreter
Reviewing your Draft: Checklist for Each Perspective

Question Askers Perspective:
- Are the core Research Questions addressed?
- For any RQs that require comparison to a past survey, are the questions asked identically?
- For any RQs that require comparison of groups (e.g., by age), are grouping questions included?

Participant Perspective:
- Consent form checklist—clear purpose and time estimate increases participation
- Question wording issues checklist—clear questions decrease attrition
- Would a participant who disliked the event/topic complete it?
- Would a time-pressed participant complete it?
- Are any other groups of participants likely to be put off? By the stated purpose, data sharing, time commitment, etc.?

Results Interpreter Perspective:
- Considering your RQs one at a time, what are the possible outcomes in terms of survey responses?
- What action will you take if you obtain each of those outcomes?
- Is there any ambiguous case, where you wouldn’t know how to interpret the result?
Still Have Questions?

Check out the Research Advice section of our website

Visit our Office Hours