Harness the Power of your Surveys: Tips and techniques for better measurement

John Kleeman MA C.Eng MBCS
Questionmark Executive Director and Founder
The food and drink at the Questionmark conference is always excellent

- Strongly Agree
- Agree
- Slightly agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

How many errors can you spot?
The food and drink at the Questionmark conference is always excellent

- Strongly Agree
- Agree
- Slightly agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Scale unbalanced
Leading question
Acquiescence bias
7 errors!
Tests and surveys use questions for different purposes

**Tests**
- Measure participant knowledge or skill
- We know what answers we want
- Participant motivated to answer well

**Surveys**
- Measure participant attitude or recollection
- We don’t know answer we are looking for
- Participant may be disinterested
Agenda

- General issues on attitude surveys
- Specific question design tips
- Satisficing and how to stop it
- Course evaluation surveys
- Questionmark capabilities that help

- Focusing on
  - Advice based on research evidence
  - Surveys of opinion and attitude in training, learning, compliance, certification
Attitude surveys
Attitudes

- Attitude
  - “feeling or opinion about something or someone”
- Measuring attitudes matter
  - influence behavior and action
- Attitudes vary significantly by context:
  - How you ask question
  - Mindset encouraged by earlier questions
  - Social desirability (what participant thinks is expected)
  - Acquiescence bias (people like to agree)
How wording influences the answer (Schuman and Presser 1980s)

- Do you think the United States should forbid public speeches in favor of communism?
  - 39% say yes “forbid”

- Do you think the United States should allow public speeches in favor of communism?
  - 56% say no “do not allow”
How question order influences answer (Cantril, 1940s)

- Asked respondents before USA joined the second world war
- More respondents said yes about the German army in Group B than Group A

**Group A**
- Should Americans be allowed to enlist in the German army?
- Should Americans be allowed to enlist in the British army?

**Group B**
- Should Americans be allowed to enlist in the British army?
- Should Americans be allowed to enlist in the German army?
Which of these is the better question?

<table>
<thead>
<tr>
<th>Agree / Disagree</th>
<th>Construct specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>My overall health is excellent. Do you:</td>
<td>How would you rate your health overall?</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>• Excellent</td>
</tr>
<tr>
<td>Agree</td>
<td>• Very good</td>
</tr>
<tr>
<td>Neither Agree or Disagree</td>
<td>• Good</td>
</tr>
<tr>
<td>Disagree</td>
<td>• Fair</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>• Bad</td>
</tr>
<tr>
<td></td>
<td>• Very bad</td>
</tr>
</tbody>
</table>
Acquiescence bias

- Strong evidence that
  - People tend to agree more than disagree in survey questions
  - Stronger when lower social status or less well educated

- Questions that ask
  - Do you agree or disagree with …
  - Please answer yes or no …
  - Will bias the answers towards agreement / yes

- Construct specific questions do not bias and also require less mental steps to answer

- Avoid “Agree/Disagree” questions
4 steps to answer a survey question

1. Comprehend the question
2. Recall/retrieve the information
3. Make a judgement
4. Select a response
Comprehension errors

1. Comprehend the question
2. Recall/retrieve the information
3. Make a judgement
4. Select a response

- Shaw “The single biggest problem in communication is the illusion that it has taken place.”
- Problems include jargon, complex words, ambiguity & misunderstanding simple words in context
- Pre-testing questions helps
Retrieval errors

- Did the participant ever know?
- Did the participant know but forget?
- If they can recall, can does the question give a good enough cue to stimulate recall?
- How accurate is the memory?
- Is the timing mixed up?
Judgement errors

- Can be strongly influenced by earlier questions
  - Continuing line of thought from before
  - Mood set by earlier questions
- Easy to misinterpret and answer what participant thinks survey author meant to ask
- Can also be influenced by social desirability - what participant thinks ought to answer
Selecting a response errors

- Rating scales limit possible responses - can only select a response from those available
  - Scales and labels can bias or influence responses
  - Open answers might be left blank or brief, hard to analyze
- May edit answer to make socially acceptable
Question design

Can you criticize these examples?
Where did you grow up?

- City
- Town
- Farm
- Country
The Mac operating system rarely gets infected by viruses and so our organization should move to purchase more Mac computers?

- Yes
- No
How happy are you with the area you live in?

- Very satisfied
- Fairly satisfied
- Neither satisfied or dissatisfied
- Fairly unsatisfied
- Very unsatisfied
- Don’t know
How long have you been a club member?

- 1-5 years
- 5-10 years
- 10-20 years
- 20+ years
How helpful did you find the instructor and the course?

- Very helpful
- Quite helpful
- A bit helpful
- Not helpful
How would you rate on-campus accommodation?

- Satisfactory
- Good
- Excellent
A useful checklist for survey questions. Avoid:

- Double barrelled questions (asking two things)
- Poor grammar and spelling
- Leading questions
- Questions open to interpretation
- Long or complex questions
- Negatives and especially double negatives
- Jargon some respondents may not know
- Unbalanced scales
- Too many alternatives
How many options should you have in a rating scale?

Less than 4
- In general reduces reliability and validity (not enough room to distinguish)

4 or 6
- Acceptable but you may get more reliability and validity with a middle option
- Risk is that people who are not sure pick one of the middle options at random

5 or 7
- Good
- Some researchers prefer 7, especially for a scale with two opposite extremes

More than 7
- In general reduces reliability and validity (too hard to distinguish)
Net Promoter Score useful 11 point scale

Based on your recent interaction with Technical Support how likely is it that you would recommend Questionmark to a friend or colleague?

Detractors
Neutral
Promoters
Order of choices can impact answer: primacy & recency

**Primacy**
- Participant chooses the first option, or the first they agree with

**Recency**
- Participant chooses the option they most recently saw

- Which conference speaker did you find most inspiring?
  - Barnum
  - Eisenberg
  - Kleeman
  - Lay
  - McNamara
  - Nash
  - Shepherd
How to reduce primacy/recency effects

- Shuffle choices when creating question

- Other approaches
  - Make question simple
  - Use short lists
  - Use open ended questions
  - If shuffling not possible as order makes sense, have 2 or more variants randomly selected
Satisficing
Satisficing

- Rather than giving optimal answer, participant chooses a “good enough” answer
- Weak satisficing
  - Less thorough in each step 1/2/3/4
- Strong satisficing
  - Skip steps 2/3 entirely, choose arbitrary or expected response
How prevalent is satisficing?
(data from a Havard Uni study 2012)

- US university survey of 250 items with a $15 cash incentive to complete
  - Response rate 65% with ~750 responses
  - 81% did some element of satisficing
  - 36% rushed through parts of the survey too fast to be giving optimal answers
  - Amount of satisficing increased later in the survey
  - Statistically impacted validity/reliability and correlations

- Satisficing likely a big issue in survey quality
How to reduce satisficing

- Have you ever answered a survey giving good enough answers rather than optimal answers?
- Why?

- Research suggests satisficing is connected with
  - Participant motivation
  - Survey difficulty
  - Participant ability
  - Participant fatigue
Some ways to reduce satisficing

**Motivation**
- Communicate benefit of the survey
- Make participant believe survey will have positive consequences
- Ask participants to carefully deliberate
- Make sure every question’s results are actionable

**Reduce difficulty**
- Keep surveys brief
- Keep questions short and simple
- Deconstruct complex questions into littler ones
- Don’t ask about events that are difficult to remember
- Don’t ask why someone did something

**General**
- Occasionally ask participants to justify answers
- Put most important questions early on
- Avoid agreement questions
- Shuffle choices to avoid primacy effects
- Remove don’t know options
How long should a survey be?

- The longer the survey, the lower response rate and the more satisficing
- Question difficulty matters too
- Unless you really need more data, keep your surveys short

<table>
<thead>
<tr>
<th>Question count</th>
<th>Average time / question (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>3-10</td>
<td>30</td>
</tr>
<tr>
<td>11-15</td>
<td>25</td>
</tr>
<tr>
<td>16-25</td>
<td>21</td>
</tr>
<tr>
<td>26-30</td>
<td>19</td>
</tr>
</tbody>
</table>

Data from SurveyMonkey
Course evaluation surveys

Some additional advice inspired by Will Thalheimer’s book:
“Performance-focused Smile Sheets”
Theory of course evaluation surveys

Theory

- Was it worth it?
  - Level 4

- Did they get it?
  - Level 3

- Did they learn it?
  - Level 2

- Did they like it?
  - Level 1

Reality

- Learners are poor predictors of benefit from learning

Forgetting curve

![Forgetting curve graph](image)
Dangers of course evaluation surveys

Learners are poor predictors of how much they will remember

- Immediate post-course opinions can be over-confident
- Key to long term remembering is spaced repetition and retrieval practice (e.g. tests/quizzes)

Application of learning requires understanding

- Surveys can’t measure this
- Give a test or quiz to measure understanding

Motivation and supervisor support are critical for applying learning

- Can be but often not asked in course evaluation surveys
- Sensible to ask questions on these
Tips on course evaluation surveys

1. Use tests or quizzes to measure knowledge and understanding
2. Avoid Likert Scale (agree/disagree) questions
3. Avoid questions with bias e.g. “My skills and abilities improved as a result of the training”
4. Consider delayed surveys (e.g. 2-4 weeks after training) to measure how training actually helped in the workplace
5. Useful to measure motivation, application and after-training support
6. Only use questions that can lead to actionable results
I learned new knowledge and skills from the training course

A. Strongly Disagree
B. Disagree
C. Slightly Disagree
D. Neutral
E. Slightly Agree
F. Agree
G. Strongly Agree
In regard to the concepts taught in the course, how motivated will you be to use these skills in your work?

A. I will NOT MAKE THIS A PRIORITY when I get back to my day-to-day job
B. I will make this a PRIORITY – BUT A LOW PRIORITY when I get back to my day-to-day job
C. I will make this a MODERATE PRIORITY when I get back to my day-to-day job
D. I will make this a HIGH PRIORITY when I get back to my day-to-day job
E. I will make this ONE OF MY HIGHEST PRIORITIES when I get back to my day-to-day job
In regard to the course topics taught, HOW ABLE ARE YOU to put what you've learned into practice on the job?

A. I'm NOT AT ALL ABLE to put the concepts into practice

B. I have GENERAL AWARENESS of the concepts taught, but I will need more training / practice / guidance / experience TO DO ACTUAL JOB TASKS using the concepts taught

C. I am ABLE TO WORK ON ACTUAL JOB TASKS, but I'LL NEED MORE HANDS-ON EXPERIENCE to be fully competent in using the concepts taught.

D. I am ABLE TO PERFORM ACTUAL JOB TASKS at a FULLY COMPETENT LEVEL in using the concepts taught.

E. I am ABLE TO PERFORM ACTUAL JOB TASKS at an EXPERT LEVEL in using the concepts taught
Selected Questionmark capabilities that help
Flexible and easy item authoring for surveys

- ~20 question types
- Item bank, so easy to re-use
Responsive Design means effective mobile delivery

One Assessment...

...many devices and browsers
Lots of ways of reporting on surveys

Three Reporting Solutions

- Reporter
- Analytics
- Results API
Anonymous surveys

**How to do it**
- Mark anonymous when authoring
- Schedule survey but system won’t record names
- Check documentation re demographics/special fields

**Why useful**
- Anonymity will give more accuracy with “sensitive” questions
- Reduce social desirability bias
- Useful for employee attitude surveys
Randomize questions / choices

**How to do it**

- Standard capability in authoring
  - Shuffle choice order
  - Randomize question order
  - Randomize questions

**Why useful**

- Shuffling choices avoid primacy/recency effects
- Randomizing question order reduces order-based bias
- Randomizing question selection can:
  - Select equivalent questions to avoid bias
  - A/B testing
  - Reduce length of survey
Jumps

How to do it

- Standard capability in authoring to jump based on
  - Score
  - Outcome to a question (most useful in surveys)

Why useful

- Direct participant to appropriate questions based on earlier answers
- Reducing apparent length of survey will reduce satisficing
Allow participants to comment on questions

**How to do it**

- When authoring, assign comment box to question

**Why useful**

- In pre-testing a survey useful for participant to explain their answer or issues on the question
- Could be used to get justification of answers to get qualitative input and reduce satisficing
Reprise – some key takeaways

- Stop using Agree / Disagree questions
- Reduce satisficing: keep survey short/simple & motivate participants
- Train your team to avoid common errors including leading questions
- Because of possible errors, comparisons or trends from previous iterations invaluable
- If your survey is important, pre-test it

Slide 51
## Recommended resources

### Books
- **Handbook of Survey Research**
  - by Marsden/Wright
- **Web Survey Methodology**
  - by Callegaro/others
- **Performance Focused Style Sheets**
  - by Will Thalheimer

### Web
- Coursera online course “Questionnaire Design for Social Surveys”
- Good article with bias examples: [http://www.cdc.gov/pcd/issues/2005/jan/04_0050.htm](http://www.cdc.gov/pcd/issues/2005/jan/04_0050.htm)
- Questionmark Blog
And now the test! How would you reword these to make them more effective?

1. Is your work made more difficult because you are expecting a baby?
2. Should people be allowed to protect themselves from harm by using XXX as self-defense?
3. What do you like about tablet computers?
4. How often do you exercise:
   - Not at all
   - Occasionally
   - Regularly
5. In the last 12 months, have you had an accident causing head injury?
Your questions