

Writing questions: some basic considerations

Patten Smith

Director of Survey Methods, Ipsos MORI

Introduction

In this document I have put together a list of 22 recommendations for practising questionnaire writers which I have drawn from personal experience of writing questions, from the question design literature and from consultation with knowledgeable colleagues. The list is certainly not exhaustive, contains items that are not completely mutually exclusive, and includes a number of recommendations that are in tension with one another. Every recommendation is, however, based upon at least some empirical evidence.

Recommendations are listed below under the following headings, but in no particular order under each heading:

- A. things you *should* do;
- B. things you *avoid* doing;
- C. things you should *give some informed thought to* before deciding what to do.

Summary of recommendations

A. Things you should do

1. Keep questions simple and easy
 - a) As a rule, keep questions short
 - b) Keep the task simple
 - c) Use familiar words and concepts
 - d) Use short showcard and read-out lists
2. Label all points on attitude scales
3. Use between five and seven points on attitude scales
4. Put open questions before closed ones on a similar topic
5. Ask general attitude questions before you ask specific ones (usually)
6. If, in an interviewer administered questionnaire, you have lots of similar filter questions ask them in a bank before asking follow-up questions
7. Start the questionnaire with easy questions they can all answer that are relevant to what respondents have been told the survey is about
8. Ask sensitive and demographic questions at the end of the questionnaire
9. For tracking questions hold wording and context constant

B. Things you avoid doing;

10. Avoid ambiguity
11. Avoid leading questions
12. Avoid double barreled questions
13. Avoid double-negatives (usually with agree-disagree questions):
14. Don't ask people to remember too much
15. Do not put the most popular items at the top of showcard lists
16. Avoid including code numbers on showcards
17. Don't ask people to rank too many items

C. Things you should give informed thought to

18. Think carefully before using open questions
19. Be alert to question sensitivity
20. Be aware of acquiescence bias with agree-disagree questions
21. Think about whether or not to use middle alternatives in attitude scales
22. Think about whether or not to use don't know codes in attitude scales

A. Things you should do

1. Keep questions simple and easy

a) As a rule, keep questions short - early writers suggested 20 words or so

But note that there is evidence that including some redundancy in *questions where you are asking people to remember things* helps people remember (perhaps because it gives respondents more thinking time). So, in these cases somewhat longer questions may be better than short ones.

b) Keep the task simple

Do not ask respondents to do much cognitive work in answering your questions, because a good many won't do it. In particular, do not include lots of information you expect them to remember in the question text, and do not ask them to perform mental tasks like working out means or percentages!

See 1c) and 10 below for examples of questions containing too much information.

c) Use familiar words and concepts

Avoid "academic" language: "main" not "principal"; "job-related" not "vocational"; "do you think that the law will affect what you have to do?", not "do you consider that the Act is likely to apply in your case?"

Also use concepts in manner in which they are *understood* rather than in technical sense: eg "percentage", "proportion", "average".

Avoid unfamiliar areas of knowledge: many will not know if they have an *endowment* or *repayment* mortgage, what a *triage nurse* is, or what the *Crown Prosecution Service* is (all examples from real surveys)

If you want to ask people about something they do not know about, don't give them an essay in an effort to educate them. The following is the approach used on the early UK lifelong learning surveys. Its obvious inadequacies were amply demonstrated by subsequent cognitive testing.

"Learning can mean practising, studying or reading about something. It can also mean being taught, instructed or coached. This is so you can develop skills, knowledge, abilities or understanding of something. Learning can also be called education or training. You can do it regularly (each day or month) or you can do it for a short period of time. It can be full time or part time, done at home, at work, or in another place like a college. Learning does not have to lead to a qualification. We are interested in any learning you have done, whether or not it was finished.

Turning to learning in general: which one of these statements most applies to you?

I am currently doing some learning activity

I have done some learning activity in the last three years

I have studied or learned but it was over three years ago

I have not studied or learned since I left full-time education"

d) Use short showcard and read-out lists

There is considerable evidence that people don't read to the bottom of long lists on show cards. Hence items at the top of the list are more frequently chosen than those further down (so-called "primacy effects").

To deal with this, keep show card lists short (10 items at most – preferably fewer). If you have more items than 10, you should either have the interviewer read out the items individually or you should use shuffle cards.

It is also advisable to reverse or rotate the placing of items on showcard list in order to minimise the net effect of any remaining primacy effects.

Generally speaking, important information about the incidence of behaviours / events should not be collected by including these behaviours /events on showcard lists, as this will lead to under-reporting. There is really no substitute for asking the question directly, as in: "In the last 12 months have you done X? Yes/No".

To give a close to home example, the following Ipsos MORI question relates to behaviour and in any case to is too long for the showcard administration used; it would have been better asked using either a read-out format or using shuffle cards:

“SHOWCARD X Which, if any, of these do you or members of your family use or benefit from?

- Council tax collection and enquiries
- Policing services
- Recycling
- Public car parks
- Hospitals
- GPs
- Dental services
- Council Housing
- Special schools
- Primary schools
- Secondary schools
- Facilities for young people
- Evening classes
- Leisure centres and swimming pools
- Libraries
- Parks, playgrounds and open spaces
- Sports facilities in parks
- Country parks
- Sports Courses and activities
- Arts and events
- Museums
- Play schemes/holiday activities
- Social services – children and families
- Social services – adults and older people
- Planning
- Trading standards and consumer advice
- Public transport

"How satisfied are you with the amount of pay you receive in your current job?"

"How satisfied are you with the kind of work you have to do in your current job?"

etc

Then ask:

"Now taking all these things into consideration, how satisfied are you with your current job?"

This approach will, of course, only work if you cover *all* important aspects of the job in the specific questions.

6. If, in an interviewer administered questionnaire, you have lots of similar filter questions ask them in a bank before asking follow-up questions

A good example is provided by the British Crime Survey where respondents are asked a long series of filter questions in the form of:

"During the 12 months since DATE have you had anything you were carrying stolen - out of your hands or from your pockets or from a bag or case?"

"[Apart from anything you have already mentioned], in that time has anyone TRIED to STEAL something you were carrying out of your hands or from your pockets or from a bag or case?"

"And [apart from anything you have already mentioned], in that time has anything (else) of yours been STOLEN, from a cloakroom, an office, a car or anywhere else you left it?"

etc

Each "yes" answer is followed by long series of questions on when, where, what happened and what respondent did about it. If each set of follow-up questions was asked immediately after its respective filter, respondents wanting to get the interview finished quickly might become tempted to say "no" to later filter questions simply to speed things along.

7. Start the questionnaire with easy questions they can all answer that are relevant to what respondents have been told the survey is about

Otherwise respondents may feel they have been interviewed under false pretences.

8. Ask sensitive and demographic questions at the end of the questionnaire

This gives time gives time to build up rapport before sensitive questions are asked. Also if respondents do not like them and decide to stop the interview, most of the data will already have been collected.

Put demographic questions at end because some (eg income) are sensitive and because they may not be seen as relevant to main topic, so it is a good idea to allow rapport to be built up first.

If you have to put sensitive or demographic questions near the beginning of the questionnaire for filtering purpose, it is important that you *explain why you are asking them*.

9. For tracking questions hold wording and context constant

It is hopefully obvious that if you want to look at how particular answers to a question change over time you should hold the working of that question constant!

Perhaps slightly less obvious is the prescription to hold constant immediately preceding questions. Context effects (ie effects of preceding questions on answers) are often not predictable, and it is therefore safest to embed your key question in the same question set every time you ask it.

B. Things you avoid doing;

10. Avoid ambiguity

Common words are often ambiguous/vague: eg "family" (how wide?), "neighbourhood" (how extensive?), "you" (in a household survey, does this mean you personally or your household?), "have" (own, have use of...?), "room" (includes box-rooms, toilets, utility rooms, etc?), "regularly" (how often does something have to be to be counted as being done "regularly"?), "in the last year" (in the last 12 months, or in the last calendar year)

Sometimes ambiguous or vague terms should be *defined* – eg "by regularly I mean at least once a week". But note there is a tension between this and the counsel for simplicity discussed above: ONS perhaps overdid it with:

"I want to ask you about all the rooms you have in your accommodation, that is including any rooms you sublet to other people and any rooms you share with people who are not in your household. Including toilets, kitchens and bathrooms, how many rooms do you have in your accommodation?"

Note that ambiguity can also arise because of how questions *sound*. For example: "were you out of the UK for **four nights (fortnight)** or more for any of these business trips?"; "Do you ever buy **mints (mince)**?" Therefore you should *read your questions out loud* before you finalise them.

11. Avoid leading questions

Questions should not give the impression that one answer is more appropriate another. See for example:

"Have you heard of the famous writer John Vooslen?" (16% yes)

"Do you agree with Nelson Mandela's /Ian Huntley's views on?"

"How often do you take a foreign holiday?" (the respondent may not)

"Do you prefer to be examined by a doctor of your own sex?" (compare: "would you rather be examined by a male or by a female doctor or doesn't it matter which?")

12. Avoid double barreled questions

These are questions that ask two or more questions at once. For example

"Do you think that the budget changes in state benefits will make people on low incomes, whether pensioners or of working age, better or worse off than they were?"

(Compare:

"Do you think that the budget changes in state benefits will make people of working age on low incomes better or worse off than they were?"

"Do you think that the budget changes in state benefits will make pensioners on low incomes better or worse off than they were?"

"Please tell me whether you agree or disagree with the following statement: we trust too much in science and not enough in religious faith"

(Compare:

"We place too much trust in science"

"We should place more trust in religious faith")

13. Avoid double-negatives (usually with agree-disagree questions):

"Please tell me whether you agree or disagree with the following statement: employers should not make special allowances for people with disabilities"

Some people have difficulty understanding what it means to *disagree* with employers *not* making special allowances

14. Don't ask people to remember too much

If events are not salient, they are hard to remember. It is also hard to remember particular instances when many similar events took place (eg a particular use of an ATM machine)

For hard-to-remember events use shorter reference periods (eg one month instead of six; a week instead of a month).

Can also put memory cues in question in the form of examples, for example:

"Do you belong to any organisations – *for example unions, religious organisations, clubs....*, (etc)?"

But the list of examples needs to be exhaustive, or there is a danger that the question will become leading.

Another approach is to use a card with examples on it to aid recall.

There is also evidence that first asking respondents to complete a calendar on which they mark important life events can help recall and date events

15. Do not put the most popular items at the top of showcard lists

There is evidence that some respondents "satisfice"¹ when answering showcard questions, reading only enough items to allow them to give what they regard as an acceptable answer. It would seem likely that satisficing respondents will be able to get away with reading fewer items if frequently

¹ A term introduced to survey research by Jon Krosnick, referring to a tendency for some respondents to do as little mental work as they can get away with when answering survey questions.

endorsed items are concentrated near the top of the list; Bobby Duffy has presented evidence suggesting that this does indeed happen. It is therefore better to distribute the frequently endorsed items throughout using a reversed card for a random half-sample, or, arguably to concentrate frequently items at the bottom of the list and *not* to use a reversed card.

16. Avoid including code numbers on showcards

There is evidence that the numbers used affect the way respondents choose their answers. That is use:

Very good

Fairly good

Fairly bad

Very bad

and not:

1 Very good

2 Fairly good

3 Fairly bad

4 Very bad

17. Don't ask people to rank too many items

Sometimes we ask people to rank items in order of importance. For example:

"SHOW CARD Please look at this card and tell me which one thing on the list you would most prefer in a job? RECORD ANSWER Which comes next? Which is third most important? And the fourth most important?"

High income

No danger of being fired

Short working hours

Opportunities for promotion

Work that gives a feeling of accomplishment"

As a rule of thumb this becomes difficult for people to do with more than 7-8 items. If you have a longer list of items, ask respondents only for 1st, 2nd and 3rd most importance.

C. Things you should give informed thought to

18. Think carefully before using open questions

Problems with open questions:

- they are dependent on interviewer probing skills

- they give some answers that are impossible to code
- they favour the articulate
- they are expensive to code
- coding produces error (often significant)
- they are usually answered badly on self completion questionnaires.

When to use them:

- if don't know full range of relevant response categories
- if you want to collect quotes
- sometimes for sensitive questions (some evidence that they get better answers to these)

19. Be alert to question sensitivity

Questions asking people to report on socially "undesirable" or private behaviours (eg drug use, sexual activity, getting drunk, being convicted of an offence, income) are usually sensitive and will lead to underreporting.

The following approaches have all been found to be helpful when asking sensitive questions at one time or another one.

- Emphasise confidentiality of answers
- Explain why it is important to collect this information
- Use a self-completion methodology (eg CASI). This has been show to be especially effective.
- Use open questions (eg with "On the days when you drink alcohol, how many drinks do you usually have - would you say one, two or three or more?"; there is evidence that heavier drinkers are reluctant to place answers at top of scale because this makes them feel their behaviour is more extreme)
- Avoid unnecessary detail. People are sometimes more willing to give an approximate indication rather than an exact answer (eg rather than ask for exact salary, ask: "is your annual gross salary less than £20,000, between £20,000 and £40,000 or over £40,000?")
- Ask respondents to pick letters from a card, rather than report the answer directly. It is probably best to associate letters with codes randomly rather than have the letters running alphabetically down the card. For example:

"The next question in on income. We want to know if this influences people's experiences of crime. Choose the letter from this card which represents the group in which you would place your total household income from all sources before tax and other deductions?"

- G. Under £2,500
- C. £2,500-£4,999
- E. £5,000-£9,999

etc"

- Ask “ever” before asking about the behaviour you want to know about; for example:
 1. Did you ever, even once, take something from a store without paying for it?
 2. And have you done this in the past 12 months?
- Try question loading techniques. A range of loading techniques have been suggested over the years. Amongst these, the following have been demonstrated to have an effect:
 - presenting valid reasons for behaviours in order to minimise the sense that they are negatively valued: (eg "in any election, some people are not able to vote because they are sick or busy or have some other reason and others do not want to vote. Did you vote in the election held on DATE?")
 - claiming that everybody does it (“as you know, many people have been killing their wives these days. Do you happen to have killed yours?”)
 - assuming the behaviour (ask “how many cigarettes do you smoke each day?” without a preceding filter question)
 - invoking authority (“many doctors now think that killing wives is good for the health. Have you ever killed your wife?”)

20. Be aware of acquiescence bias with agree-disagree questions

Acquiescence bias is the tendency of people to be more likely to agree with a statement than to disagree with it.

For example in one survey 10% of respondents agreed with *both* the following statements:

"It is hardly fair to bring children into the world with the way things look for the future"

"Children born today have a wonderful future to look forward to"

Acquiescence bias is greater for less well educated respondents. As a result of this spurious correlations with agree-disagree statements may arise.

In an effort to get around this problem, researchers often use multiple items where “agree” denotes both positive and negative attitudes, but doing this means that those who acquiesce more will be more likely to get middle scores. Also, it is often hard to write a negative version of a positive statement or vice versa without introducing a double negative (see 13 above).

Overall then, you should try to avoid writing questions which use an agree-disagree format unless you have compelling reasons to do so (indeed, some methodologists, notably Jon Krosnick, argue that we shouldn't use agree-disagree scales at all, and claim that it is always possible to find a better alternative).

That said, there will still be some occasions where it is necessary to use the agree-disagree format – notably when you are collecting tracking data based on this format and when you wish to construct a formal Likert scale.

21. Think about whether or not to use middle alternatives in attitude scales

Should one include a middle “neutral” point on a bipolar intensity scale (eg include “neither agree nor disagree” option in agree-disagree scale)? There is much dispute: about this. One argument is

that many giving neutral response still tend to lean one way or the other, and we want to pick this up. The other argument is that we should allow for people who have a genuine neutral opinion.

The research indicates that ratio of the number of respondents agreeing to the number disagreeing is usually similar whichever format used. Also, relationships with other variables are usually similar whichever format is used.

So, there is no single right answer. It depends if your priority is to give a voice to those with a genuinely neutral position, or to push respondents in one or another direction

22. Think about whether or not to use don't know codes in attitude scales

For subjects where people may not have an opinion, there is evidence that some people with no opinion effectively answer on a random basis (Converse coined the phrase “non-attitudes” to describe this). It may, therefore, be advisable offer an explicit “don't know” option in the question or on the show card.

However, it is also worth noting that Krosnick has argued that some “lazy” respondents, who would express an opinion if pushed, will use the "don't know" code to save themselves mental effort.

You will therefore need to decide on a case-by-case basis whether to include an explicit "don't know" code taking into account the respondents' likely familiarity with the question topic.