SSOAR

Open Access Repository
www.ssoar.info

Question Wording (Version 2.0)

Lenzner, Timo; Menold, Natalja

Erstveröffentlichung / Primary Publication
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Lenzner, T., & Menold, N. (2016). *Question Wording (Version 2.0).* (GESIS Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/gesis-sg_en_017

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

Terms of use:
This document is made available under a CC BY-NC-ND Licence (Attribution-Non Comercial-NoDerivatives). For more Information see:
https://creativecommons.org/licenses/by-nc-nd/4.0

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

gesis Leibniz-Institut für Sozialwissenschaften

GESIS Survey Guidelines

Question Wording

Timo Lenzner & Natalja Menold

December 2016, Version 2.0

Abstract

This contribution provides an overview of the basic rules that should be observed when formulating survey questions. It is structured along the cognitive response process that respondents undergo when answering survey questions. This process comprises a total of four tasks: understanding the question text, retrieving information from memory, forming a judgment, and formulating a response in the format provided. Using practical examples, the authors demonstrate how questions can be worded in such a way that it makes it easier for respondents to carry out these four tasks.

Citation

Lenzner, T., & Menold, N. (2016). Question Wording. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_017

1. What requirements must the wording of survey questions meet?

A survey question should be interpreted by all respondents in the manner intended by the questionnaire designer and should be as easy as possible to understand and answer. This is a prerequisite for error-free data collection. Question wording plays a decisive role in this regard because it directly influences the extent to which respondents find it easy or difficult to understand the question in the manner intended by the researcher and to answer it.

When answering a survey question, respondents must carry out four cognitive tasks (Tourangeau, 1984; Tourangeau, Rips, & Rasinski, 2000):

(1) understanding the question and drawing inferences about the point of the question (*comprehension*);

(2) retrieving from memory relevant information to answer the question (*retrieval*);

(3) making a judgment on the basis of this information (*judgment*);

(4) reporting this judgment in the appropriate response format (*response*).

Depending on the way in which a survey question is formulated, each of these tasks may be more or less difficult to perform. For example, vague terms or a complex sentence structure may render comprehension of a question difficult (Tourangeau et al., 2000). When respondents are asked about events that happened a long time ago, they usually find it difficult (a) to retrieve from memory the relevant information to answer the question and/or (b) to form a judgment on the basis of that information (Cannell, Miller, & Oksenberg, 1981). Response categories that are not unequivocal and mutually exclusive make it difficult for the respondents to express their responses in the appropriate response format. Sub-optimal survey question wording may therefore make questions difficult to answer and may impair the accuracy of the responses. For example, comprehension problems may distort the responses to such an extent that they do not reflect the reality that respondents are being asked to describe.

The present contribution addresses how survey questions should be worded so that they keep the cognitive effort involved in answering them as low as possible and facilitate the most error-free data collection possible. It is structured along the four stages of the cognitive response process.

2. How can the comprehensibility of survey questions be improved?

In survey research, it is generally agreed that questions should be worded in such a way that they can be easily understood by the respondents, and that all respondents also understand them in the manner intended by the researcher (Fowler, 1992). To achieve these two objectives, different aspects of question wording should be taken into account.

2.1 Avoid unfamiliar or uncommon terms

Unfamiliar or uncommon terms carry the risk that the respondents may not understand them at all or may understand them differently. When respondents do not understand a term, it is hardly possible for them to meaningfully answer the question. And if two or more respondents interpret the question

differently, their responses may differ not because their opinions differ, but because they understand the question differently.

Unfamiliar or uncommon terms, such as foreign words, technical terms, abbreviations, acronyms, and rare terms, should therefore be avoided in survey questions (Graesser et al., 2006; Lenzner, 2011). It is recommended that all words in survey questions should be written out in full, and that acronyms should be avoided (e.g., Transatlantic Trade and Investment Partnership rather than TTIP, and Renewable Energy Sources Act rather than RES Act). Moreover, an (online) thesaurus should be consulted to check whether more common, and therefore more understandable, synonyms exist for central terms in survey questions. If this is the case, the less common terms should be replaced by these synonyms. The following are two examples:

(1) "To what extent do you agree with the following statement?
The social *discrepancies* (better: *differences*) in Germany will continue to exist."

(2) "How often in the last four weeks have you had *somatic* (better: *physical*) health problems?"

If more common synonyms for uncommon terms (see 3a) cannot be found, the uncommon terms should be defined (see 3b). The following example is taken from Porst (2008; our translation):

(3a) "At what age do you think *andropause* begins in men?"

(3b) "The term *andropause* describes age-related hormonal changes in men that may affect their emotions and their sex life. This process is comparable to menopause – i.e., the so-called "change of life" – in women. At what age do you think andropause normally begins in men?"

What terms and wording are easily and clearly understandable depends, of course, on the target group. For example, in a survey of physicians about the way they deal with patients with a particular illness, it would be laborious and uneconomical to avoid certain terms in the interests of greater comprehensibility (Dillman, Smyth, & Christian, 2009). In such a case, specialised terms may be used or one should check whether experts are familiar with the terms in question.

2.2 Avoid imprecise relative terms

Imprecise relative terms, such as "often," "recently," or "considerably," are problematic insofar as they implicitly refer to an underlying continuum but their exact position on this continuum is not clear (Graesser et al., 2006; Lenzner, 2011). They may therefore be interpreted in different ways, which increases the risk that respondents will not understand them in the manner intended by the questionnaire designer. For example:

(4a) "Have you visited any doctors *recently*?"
If respondents answer "yes," they are asked a further question about the number of doctor's visits.

In this question, it can be assumed that the interpretation of the term *recently* will differ significantly among respondents. Some respondents will take it to mean the last few days; others will think of the last few months. Besides the problem that different respondents interpret the term differently, there is also the risk that respondents will notice the ambiguity of the term and that this will make it more difficult for them to answer the question (Lenzner, 2012). For example, respondents may wonder what time period is meant by *recently*, and if they cannot decide on a particular time period, they may leave the question unanswered. Therefore, such imprecise relative terms should be avoided in survey questions (especially in the question text), and concrete terms should be used instead. The above-mentioned example could be revised as follows:

(4b) "Have you visited any doctors in the *past four weeks*?"

If respondents answer "yes," they are asked a further question about the number of doctor's visits.

For more information on questions where the response determines whether further questions are asked (so-called filter questions), see the *GESIS Survey Guidelines* contribution "Postal Surveys" (Menold, 2016).

2.3 Avoid abstract and ambiguous terms

Abstract terms are very easy to recognise because they cover several more specific and concrete terms. The term *cultural events* in question 5a, for example, includes events such as theatre performances, operas, pop concerts, and rock festivals.

(5a) "How often do you attend *cultural events* in your free time?"

Here, some respondents may think only of events such as theatre performances and classical concerts, while others may think only (or also) of pop concerts and poetry slams. On the other hand, because of the ambiguity of the term *cultural events*, some respondents may find it hard to decide how to interpret the term and may therefore leave the question unanswered (Lenzner, Kaczmirek, & Galesic, 2011; Lenzner, 2012). To increase question comprehensibility, concrete and specific terms rather than abstract terms should therefore be used where possible. As an alternative, the collective term *cultural events* can be illustrated with three examples (Tourangeau et al., 2014). For example, question 5a could be revised as follows:

(5b) "How often do you attend *cultural events, such as concerts, readings, or theatre performances,* in your free time?"

Another way to make the question easier to answer would be to present a list of different events to the respondent for assessment.

2.4 Avoid complex sentence structures

To avoid overburdening the cognitive capacity of the respondents, the complexity of the syntactic structure of survey questions should be as low as possible. For example, the grammatical complexity of questions can be reduced by converting passive sentences (e.g., question 6a) into active sentences (e.g., question 6b; Lenzner, Kaczmirek, & Lenzner, 2010):

(6a) "To what extent do you agree with the following statement?
Too much money *is spent* by the government on assisting immigrants."

(6b) Government *spends* too much money assisting immigrants."

Moreover, the nominalisation of verbs, that is, the transformation of verbs into nouns (e.g., *acquire* → *acquisition*), should also be avoided:

(7a) "To what extent do you agree with the following statement? I have enough time during working hours for the *acquisition* of new knowledge."

(7b) "I have enough time during working hours to *acquire* new knowledge."

2.5 Avoid hypothetical questions

A hypothetical question is one where, in order to answer the question, the respondents have to put themselves in a situation that does not reflect their reality. For example:

(9) "Suppose you were the Federal Chancellor, what problem in this country would you tackle first?"

(10) "Suppose you had a grown-up daughter, would you want her to make a career?"

Hypothetical questions are problematic, especially because they require the respondents to mentally put themselves into an imaginary situation, to keep it in their short-term memory, and then to answer the question from this hypothetical situation. How well this (dual) cognitive task can be performed by respondents depends, first, on their cognitive abilities; second, on whether they have ever dealt with the hypothetical situation in the past; and third, on how far removed from reality the hypothetical situation is (Porst, 2008).

Instead of formulating questions hypothetically, respondents can be asked to assess the issues of interest in the here and now. For example, one does not have to be the Federal Chancellor to assess the priority of a problem. So respondents could be asked: "What problem should the Federal Government tackle first?"

2.6 Avoid double-barrelled questions

Double-barrelled questions are questions that address two, or even more, different matters (stimuli). Respondents find such questions difficult to answer if they evaluate the individual stimuli differently. For example, the following question (11) contains four stimuli:

(11) "To what extent do you agree with the following statement?
The State must ensure that people have a good livelihood even in case of *sickness*, *need*, *or unemployment* and in *old age*."

Respondents who are in favour of State support in all four areas do not have any difficulties answering the question. They can simply answer "Completely agree," or "Agree". However, the situation is different when respondents are in favour of State support in the case of "sickness" and "need" but not in the case of "unemployment" and in "old age". These respondents cannot answer the question with the response categories provided (e.g., their response would be "In case of sickness and need, I agree. In case of unemployment and in old age, I do not agree."), and if they do answer, it is impossible to know which of the four stimuli their response refers to.

There are different possibilities of avoiding double stimuli. For example, one can choose one alternative and limit the question to it:

(12) "The State must ensure that people also have a good livelihood in case of unemployment.

A further possibility is to have respondents assess each stimulus, or individual stimuli, in the form of rating scales or forced-choice questions (e.g., question 13; see also Section 4.2):

(13) "In which of the following cases should the State provide enough financial support to enable people to have a good livelihood?

Unemployment because of own sickness	Yes/No
Unemployment	Yes/No
Insufficient pension in old age	Yes/No"

However, researchers may not be interested in details at all, but may wish to generally determine whether respondents are in favour of State support when people have no or low income. In that case, the individual stimuli can serve as examples:

(14) "The State must ensure that people with low income have a good livelihood, for example in case of unemployment or inability to work, or in old age."

2.7 Avoid negations and double negatives

Questions with double negatives are particularly difficult to understand. They usually require the respondents to accomplish a complex logical reasoning task (Foddy, 1993; Fowler, 2001). Double negatives occur, for example, whenever respondents must respond to a negatively formulated statement by means of an agreement scale (*completely agree* – *completely disagree*). For example:

(15a) "To what extent do you agree with the following statement?
Poorer countries should *not* be expected to make less effort to protect the environment than rich countries."

In order to express a positive opinion on the subject of this statement ("Less effort to protect the environment should be expected from poorer countries.") the respondents must *dis*agree with the negatively formulated statement. To do so, they must first "translate" the negative statement into a positive statement and then decide whether they agree or disagree with it (Akiyama, Brewer, & Shoben, 1979). Positively formulated statements are easier to understand because the "translation step" can be omitted:

(15b) "To what extent do you agree with the following statement?
Poorer countries should be expected to make less effort to protect the environment than rich countries."

Negations – words like "no" or "none" – should be avoided as they can easily be overlooked (Dillman et al., 2009).

2.8 Avoid (implicit) presuppositions

Presuppositions are present whenever the question text regards as given a situation that may not necessarily apply to the respondents. Such presuppositions should be avoided in survey questions because respondents who do not share, or agree with, them cannot answer the question correctly. The following is an example of a presupposition (see Fillmore, 1999):

(16) "Family life often suffers because men concentrate too much on their work."

This statement presupposes that men concentrate too much on their work. The respondents are asked to state whether, or to what extent, this is, in their opinion, one reason why family life suffers. Respondents who are not of the opinion that men concentrate too much on their work cannot actually answer this question because neither *agree* nor *disagree* expresses their actual opinion.

Presuppositions also occur when the question does not apply to the respondents or when the content of the question does not affect the respondents (Dillman et al., 2009). For example, the question "Are you considering having your house energetically refurbished?" applies only to house owners and only to those house owners who have not yet had the energy performance of their house upgraded. In the case of a question like this, it would therefore be appropriate to ask first whether the respondent is the owner of the house and then whether the house has already been energetically refurbished. With the help of filters, the question about the intention to have the house energetically refurbished can then be addressed only to those house owners whose houses have not yet been refurbished in this way. However, when using filter questions, the mode of data collection should be taken into account. In the case of computer-assisted surveys, filtering is easily implemented and does not cause respondents any problems. In the case of postal surveys, on the other hand, numerous and complex filters should be

avoided, which generally limits the collectability of certain, and very detailed, information (see the *Gesis Survey Guidelines* contribution "Postal Surveys," Menold, 2016).

2.9 Avoid leading questions

A leading question is one that suggests or provokes a certain response. The following example is taken from Porst (2008; our translation):

(17a) "Leading scientists are of the opinion that car emissions can impair children's growth. Do you think this opinion is correct or do you think it is incorrect?"

Phrases such as "leading scientists" and "most people" often result in respondents' lacking the courage to contradict the authority to whom the statement is attributed or to contradict the majority. Such phrases may therefore provoke conforming responses (Porst, 2008). As we have often observed in cognitive pretests, leading questions may also cause respondents to ponder on the suggestion and to make less effort to understand and appropriately answer the question. This in turn would have a negative effect on the quality of their responses. Leading questions should therefore be avoided in surveys, and question 17a above can be rephrased as follows:

(17b) "Do you think that the statement that car emissions can impair children's growth is correct or incorrect?"

A special case of a leading question is where only one of the two possible answers to a dichotomous question is explicitly formulated (Faulbaum, Prüfer, & Rexroth, 2009). For example:

(18a) "Do you think that all the workers in a company should be in the union?"

In the case of such "unbalanced" questions, there is a risk that the formulated alternative will win more approval than the one that is not formulated (Schuman & Presser, 1981). To avoid such an effect in dichotomous questions, both response alternatives should be explicitly formulated:

(18b) "Do you think that all workers in a company should be in the union or that it should it be left up to each individual?"

Moreover, in question 18a, graduated responses could be collected with the help of a rating scale (see Saris & Gallhofer, 2007; Dillman et al., 2009), and an introductory question – "To what extent do you agree with the following statement?" – could be used to request respondents to assess their level of agreement with the statement.

3. How can the recall of relevant information and the formation of judgments be facilitated?

Ideally, the answer to a question is directly accessible, and all that respondents have to do is retrieve it from memory. However, this ideal case is more the exception than the rule, and it can be expected only in the case of socio-demographic questions such as "In what month and year were you born?" or "How many children under the age of 18 do you have?" (see Porst, 2008). When answering most survey questions, respondents must first search their memory for relevant information and then generate a response.

In the case of *attitude questions*, respondents must either recall and retrieve an existing opinion, or, in the absence of such an opinion, they must recall relevant information that enables them to form a

judgment immediately in the survey situation. How easy or difficult respondents find this, depends mainly on how important the topic of the question is to them, how strong their attitudes to the topic are, and how much experience they have of it in the first place. It is highly likely that marijuana users will find the question "Do you think that the use of marijuana should be made legal or not?" easier to answer than will those who have never, or hardly ever, given any thought to the legalisation of soft drugs. In this case, questionnaire designers have few possibilities of facilitating the recall of relevant information by formulating the question in a particular way.

The situation is different in the case of *behavioural questions* (e.g., "How many times did you visit a dentist in the past year?"). To answer such questions, respondents must recall relevant events, date them correctly, and count or estimate their number. The ease with which relevant information can be recalled depends mainly on the time that has elapsed since the event in question and on the significance of that event for the respondent (Cannell, Marquis, & Laurent, 1977). This has a number of implications for the formulation of survey questions.

3.1 Choose a suitable time frame for the question

The time frame of a survey question should be comparatively short when:

- the content relates to everyday things of little importance, for example: "How many times have you eaten ready meals in the past *5 days*?" (instead of *30 days*, for example);
- a frequent but irregular behaviour is asked about, for example: "How many text messages have you sent in the past *3 days*?" (instead of *4 weeks*, for example);
- a regular behaviour is asked about, for example, "How many times have you been to the hairdresser in the past *3 months*" (instead of *12 months*, for example; provided people go to the hairdresser regularly rather than rarely).

If the time frame of a question is short, the respondents simply count and report the number of events that they can recall. In the case of longer time frames, they typically count only the number of events that occurred in a short time period and calculate the number for the entire time period on this basis. Hence, using a longer time frame does not, as assumed, yield more correct information, because the information collected is based only on a shorter period.

However, there are also situations in which a longer time period should be used, namely when:

- the question relates to very significant or important events that are easily recalled and that do not occur often, for example: "How many times have you moved house in the *past 10 years*?" or "How many times *in your life* have you been married?", or when
- the question asks about a behaviour that is infrequent or irregular, for example: "How many times have you participated in a medical screening in the *past 12 months*?".

If shorter time periods were used in such behavioural questions, it is highly likely that a large number of "0" responses would be obtained.

3.2 Use unambiguous reference periods

As a general rule, reference periods should always be as unambiguous as possible. For example, the time frame of question 20a below can be clarified either as shown in question 20b or even as shown in question 20c:

(20a) "Have you visited any doctors *recently*?"

(20b) "Have you visited any doctors in the *past four weeks*?"

(20c) "Have you visited any doctors *since the 1st of April this year*?"

Stating the date makes the time frame clearest, but it does not guarantee that the respondents will find it easier to recall the relevant information. Calendar dates are not well represented in human memory, and they are not usually saved as part of events. Respondents may be able to clearly recall the situation where they bought their second-last bicycle, but they may not recall the date (see Porst, 2008).

3.3 Divide complex questions into several individual questions

Dividing complex questions into several questions may facilitate the recall of relevant information and the formation of a judgment. For example, after asking whether a person has visited a doctor in the past 12 months respondents are asked "How many different doctors have you visited in the past 12 months?" This question may be hard to answer in the case of a relatively large number of doctor's visits, and it is likely that the one or other of these visits will not be recalled. It is therefore appropriate to divide the question into several successive questions (Fowler & Cosenza, 2008):

(21a) "Have you visited any *primary care doctors/general practitioners* in the past 12 months? (If yes: How many?)"

(21b) "Have you visited any *specialists* in the past 12 months? (If yes: How many?)"

(21c) "Have you visited any *homoeopaths* in the past 12 months? (If yes: How many?)"

(21d) "How many *other doctors* whom you have not mentioned have you visited in the past 12 months?"

Such a procedure makes it easier for the respondents to recall the relevant information and to form a judgment. First, they must focus only on one sub-area of the relevant information (in the present example, on a certain type of doctor) rather than having to focus on several areas at the same time. Second, they have more time to think about the individual questions, which makes it easier for them to recall the information. And finally, the previous questions have a certain priming function: The fact that similar information (e.g., about specialists) has already been recalled renders it easier to recall the information (e.g., about homoeopaths) that is being sought.

4. How can it be made easier for respondents to fit their "internally" determined response into the prescribed response format?

The respondents' final task when answering a survey question is to communicate the "internally" determined response in the format prescribed by the questionnaire (or by the interviewer). When doing so, the choice and design of the response format can strongly affect the way the question is answered (Couper, Tourangeau, Conrad, & Crawford, 2004; Tourangeau, Couper, & Conrad, 2004, 2007).

In general, three large groups of response formats can be distinguished:

1. Open-ended questions
   These questions include only the question text and, perhaps, instructions for completion, but they do not include any response categories. Hence, respondents answer open-ended questions in their own words (see the *GESIS Survey Guidelines* contribution "Open-Ended Questions," Züll, 2016)

2. Closed-ended questions with rating scales
   These questions include a pre-defined number of response options, which are ordered according to an inherent logic and refer to a certain dimension, for example the degree of strength or the intensity ("very strong – strong – medium – weak – not at all"). Respondents use a rating scale to express their location on the response dimension in relation to a certain characteristic (e.g., interest in politics; see the *GESIS Survey Guidelines* contribution "Designing Rating Scales in Questionnaires," Menold & Bogner, 2016).

3. Closed-ended questions without rating scales
   These questions also contain a pre-defined number of response options but not in the form of a rating scale but merely in the form of a list. The response options may be ordered according to an inherent logic. For example, in the case of a question about the highest general education school leaving qualification, the response options are ordered from the lowest to the highest possible school leaving qualification. Alternatively, the ordering of the response options may be left up to the questionnaire designer. In the case of the "Sunday" question about party preference ("If a federal election were held next Sunday, which party would you vote for?"), for example, the options may be in this order: "CDU, SPD, Greens, FDP..." or in this order: "SPD, CDU, FDP, Greens ...".

Because open-ended questions and closed-ended questions with rating scales are dealt with in detail in other contributions to the *GESIS Survey Guidelines*, only general advice on formulating response categories in closed-ended questions with ordered or unordered lists will be given here.

4.1 The response categories should match the question

Occasionally, the response options do not match the actual question. For example:

(22) "For each of the following organisations, please state whether you have been an active member of this organisation during the past 12 months.

a) Sports club
b) Trade union
c) Political party

Response categories: "Often – Sometimes – Rarely – Never"

The question text relates to *whether* the person has been active in one of the organisations in the past 12 months. In other words, the question is asked in such a way that the only possible answers are "Yes" or "No". However, the response categories require respondents to report a frequency. Because this contradiction may irritate the respondents, the question text should be aligned to the response categories ("How often were you ...?") or the response categories should be aligned to the question text ("Yes," "No").

4.2 The response categories should be exhaustive and mutually exclusive

In the case of closed-ended questions, the pre-defined response options must cover the entire spectrum of possible responses. If substantively logical response options are missing, some respondents will either not answer the question at all or will randomly choose one of the response categories (Faulbaum et al., 2009). Question 23 below lists only a few areas that respondents might consider to be important tasks of the Federal Government.

(23) "The following question relates to the recently formed Federal Government in Germany. Which political tasks do you think the new Federal Government should focus on in the next four years? Please mark all that apply!

- Development aid for poor countries
- Agricultural subsidies
- Defence, in particular the withdrawal of the armed forces from Afghanistan
- Financial crisis in the European Union"

In this question, the list should be supplemented with further important aspects – for example, the promotion of women, the phasing out of nuclear power, financial support for families with children, and increasing the number of nursery places. Moreover, as it is hardly possible to list all possible options in this case, the list should be supplemented with an "Other" category.

Questions that list a certain number of alternatives and allow respondents to choose either one alternative (so-called single-response questions) or several alternatives (so-called multiple-response questions) are very common in surveys. In the case of such response lists, response-order effects may occur. For example, if the response options are visually presented, the first options on the list may be chosen more frequently than the subsequent options (primacy effect). In the case of orally presented options, on the other hand, the last-mentioned options are chosen more often (recency effect). To avoid such response-order effects, respondents should be asked to provide an answer (yes/no) to each item on the list (so-called forced-choice format, see Smyth et al., 2006). This design was realised in question 22 above.

Moreover, in the case of closed-ended questions, it must be ensured that the response categories are mutually exclusive so that the respondents can unequivocally assign their responses to a particular category. In question 24 below, for example, respondents who get up at 6:30 a.m. can select both the second and the third response category:

(24) "What time do you usually get up on a working day?

- Before 6:00 a.m.
- Between 6:00 a.m. and 6:30 a.m.
- Between 6:30 a.m. and 7:00 a.m.
- Between 7:00 a.m. and 7:30 a.m.
- Between 7:30 a.m. and 8:00 a.m.
- After 8:00 a.m."

The response options should therefore be reformulated to render them mutually exclusive:

(25) "What time do you usually get up on a working day?

- Before 6:00 a.m.
- Between 6:00 a.m. and before 6:30 a.m.
- Between 6:30 a.m. and before 7:00 a.m.
- ...

4.3 The response categories should be logically/systematically ordered

To make the response spectrum easier for respondents to understand, response categories that are subject to a certain gradation logic should be ordered according to this logic. The order in which the response options appear in question 26 below is not substantively logical, for example:

(26) "Since you started taking the medication, have your symptoms ...

- disappeared completely,
- improved slightly,
- improved significantly, or
- not changed at all?"

The response categories "improved slightly" and "improved significantly" would have to be switched in order to produce a logical gradation sequence.

4.4 The response categories should not presuppose regularity

Some behavioural questions or factual questions presuppose that the behaviour or the events in question are subject to a certain regularity. However, this may not apply in the case of all respondents, which makes it harder for these respondents to answer the question (Fowler & Cosenza, 2008). For example:

(27a) "How often do you withdraw money from an ATM?
- Never
- Less than once a month
- Once a month
- Every two weeks
- Every week"

Question 27a, for example, presupposes that respondents withdraw cash from ATMs at regular intervals. However, this may not apply to all respondents. Some of them may withdraw cash from ATMs only on an as-needed basis (and at irregular intervals). It would therefore be better to ask about the number of times that the respondent went to an ATM within a concrete time frame. Question 27b, for example, does not presuppose a regular behaviour:

(27b) "How often have you withdrawn money from an ATM in the past month?
- Never
- Once
- Two or three times
- Four times or more"

5. Summary

As the above presentation clearly shows, a wide range of methodological aspects must be taken into account when formulating survey questions. Taking these aspects into account is a basic prerequisite for ensuring that survey questions are (a) interpreted by respondents in the manner intended by the questionnaire designer and (b) as easy as possible to understand and answer. However, uncertainty with regard to the comprehensibility and answerability of survey questions often persists even when all the formal rules discussed above have apparently been observed. Greater certainty as to whether a question in a questionnaire works as intended and is easy to understand and answer is usually achieved only by conducting a cognitive pretest (see the *GESIS Survey Guidelines* contribution "Cognitive Pretesting," Lenzner, Neuert, & Otto, 2016). However, even with the help of a cognitive pretest, it cannot be ensured with certainty that a question will reliably measure (reliability) the intended content (validity). Further theoretical and empirical steps are therefore necessary to investigate the reliability and validity

of survey questions. Readers are referred in this connection to the "Quality Assurance" section of this website and to the quality standards developed on behalf of the German Data Forum (RatSWD; Rammstedt et al., 2015).

References

Akiyama, M. M., Brewer, W. F., & Shoben, E. J. (1979). The yes-no question answering system and statement verification. *Journal of Verbal Memory and Verbal Behavior, 18*, 365-380.

Cannell, C. F., Marquis, K. H., & Laurent, A. (1977). A summary of studies of interviewing methodology: 1959-1970. *Vital and Health Statistics, 69*, i-78.

Cannell, C., Miller, P., & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Ed.), *Sociological methodology 1981* (pp. 389-437). San Francisco: Jossey-Bass.

Couper, M. P., Tourangeau, R., Conrad, F. G., & Crawford, S. D. (2004). What they see is what we get. *Social Science Computer Review, 22*, 111–127.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys. The tailored design method.* Wiley: New Jersey.

Faulbaum, F., Prüfer, P., & Rexroth, M. (2009). *Was ist eine gute Frage? Die systematische Evaluation der Fragenqualität.* Wiesbaden: VS Verlag.

Fillmore, C. J. (1999). A linguistic look at survey research. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), *Cognition and survey research* (pp. 183-198). New York: Wiley.

Foddy, W. (1993). *Constructing questions for interviews and questionnaires: theory and practice in social research.* Cambridge: Cambridge University Press.

Fowler, F. J. (1992). How unclear terms affect survey data. *Public Opinion Quarterly, 56*, 218-231.

Fowler, F. J. (2001). Why it is easy to write bad questions. *ZUMA-Nachrichten, 48*, 49-66.

Fowler, F. J., & Cosenza, C. (2008). Writing effective questions. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 136-160). New York: Lawrence Erlbaum.

Graesser, A. C., Cai, Z., Louwerse, M. M., & Daniel, F. (2006). Question understanding aid (QUAID). A web facility that tests question comprehensibility. *Public Opinion Quarterly, 70*, 3–22.

Lenzner, T. (2011). *A psycholinguistic look at survey question design and response quality.* University of Mannheim: MADOC.

Lenzner, T. (2012). Effects of survey question comprehensibility on response quality. *Field Methods, 24*, 409-428.

Lenzner, T., Kaczmirek, L., & Galesic, M. (2011). Seeing through the eyes of the respondent: An eye-tracking study on survey question comprehension. *International Journal of Public Opinion Research, 23*, 361-373.

Lenzner, T., Kaczmirek, L., & Lenzner, A. (2010). Cognitive burden of survey questions and response times: A psycholinguistic experiment. *Applied Cognitive Psychology, 24*, 1003-1020.

Lenzner, T., Neuert, C., & Otto, W. (2016). Cognitive Pretesting. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_010

Menold, N. (2016). Postal surveys. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_018

Menold, N., & Bogner, K. (2016). Design of Rating Scales in Questionnaires. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_015

Porst, R. (2008). *Fragebogen. Ein Arbeitsbuch*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Rammstedt, B., Beierlein, C., Brähler, E. Eid, M. Hartig, J. Kersting, M. Liebig, S. Lukas, J. Mayer, A.-K., Menold, N., Schupp, J. & Weichselgartner, E. (2015). Quality standards for the development, application, and evaluation of measurement instruments in social science survey research. In: *RatSWD Working Papers Series*, 245, pp. 1-35.

Saris, W. E., & Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Comparing check-all and forced choice question formats in web surveys: The Role of satisficing, depth of processing, and acquiescence in explaining differences. *Public Opinion Quarterly, 70,* 66-77.

Tourangeau, R. (1984). Cognitive science and survey methods. In T. B. Jabine, M. L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73-100). Washington, DC: National Academy Press.

Tourangeau, R., Couper, M. P., & Conrad, F. G. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly, 68*, 368–393.

Tourangeau, R., Couper, M. P., & Conrad, F. G. (2007). Color, labels, and interpretive heuristics for response scales. *Public Opinion Quarterly, 71*, 91–112.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

Tourangeau, R., Conrad, F. G., Couper, M. P., & Ye, C. (2014). The effects of providing examples in survey questions. *Public Opinion Quarterly, 78*(1), 100-125. doi: 10.1093/poq/nft083

Züll, C. (2016). Open-ended questions. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_002