

GESIS Survey Guidelines

Response Biases in Standardised Surveys

Kathrin Bogner & Uta Landrock

Abstract

Response biases are present in standardised surveys when values collected from, or reported by, respondents systematically deviate from the actual, true values. Response biases are therefore potential sources of error in social science surveys. They occur when respondents react not only to the content of the survey question – the question stimulus – but also to other factors. These factors can be found in the most diverse elements of a survey. They may lie in the respondent him- or herself, in the interview situation, or in the survey instrument. Response bias that lies in the respondent him- or herself is often a consequence of satisficing. Respondents who have a tendency to satisfice when answering questions are motivated by a desire to reduce their cognitive effort. This results in response biases such as acquiescence or moderacy bias. In the case of socially desirable responding, respondents orient themselves towards social norms in order to answer in the way in which they assume that the interviewer expects them to answer. Respondents who display extreme response bias tend to choose the extreme rating categories irrespective of the content of the question. The second source of response bias is the interview situation: Response bias also occurs when response behaviour is dependent on the situation. When respondents react to characteristics or behaviours of the interviewer, these reactions are known as interviewer effects. By contrast, the term *presence effect* is used when the respondent's response behaviour is dependent on the presence of third parties. The third source of response bias is the survey instrument: The way in which the questions in the questionnaire are formulated, and the order of the questions and the response alternatives in the questionnaire may influence response behaviour – for example as halo, primacy, or recency effects – and may lead to response biases.

Citation

Bogner, K., & Landrock, U. (2016). Response Biases in Standardised Surveys. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_016



Contents

1. Response biases in standardised surveys	1
2. Types of response biases.....	2
2.1 Socially desirable responding.....	2
2.2 Acquiescence	3
2.3 Moderacy response bias and extreme response bias	4
3. Influence of the interviewer and the interview situation	5
3.1 Interviewer effects	5
3.2 Presence of third parties.....	6
4. Influence of the questionnaire.....	7
4.1 Question-order effects	7
4.2 Response order.....	8
References	9

1. Response biases in standardised surveys

The term *response bias*, or *response set*, refers to a systematic deviation of the reported respondent values from the true values. Response bias is therefore a potential source of error in standardised surveys. The causes of this bias can be found in different elements of a survey. Certain response biases, such as socially desirable responding or acquiescence, may be a consequence of a specific response behaviour on the part of the respondent. For example, acquiescence may be a strategic response behaviour in the sense of satisficing (see below), or it may be a manifestation of a personality trait (Krosnick, 1999b). Moreover, the content of the questions, the design of the questionnaire, and characteristics of the interview situation, such as the presence of an interviewer or a third party, may contribute to the occurrence of response bias. Overall, response biases are a consequence of the fact that survey interviews are social processes in which the respondents react not only to the question stimulus but also to other aspects of the situation (Kahn & Cannell, 1957; Groves et al., 2009).

The optimal cognitive process that respondents undergo when answering a question comprises four steps (Cannell, Miller, & Oksenberg, 1981; Tourangeau, Rips, & Rasinski, 2000):

- (1) Understanding the question
- (2) Retrieving the relevant information from memory
- (3) Evaluating the completeness and relevance of the retrieved information
- (4) Giving the answer

Each individual step may involve considerable cognitive effort, and answering all the survey questions may impose a substantial cumulative cognitive burden on the respondent (see, among others, Krosnick & Fabrigar, 1997). This “high road” to information processing, which is known as optimising, contrasts with the “low road,” which is called satisficing. Respondents who engage in weak satisficing still execute all four cognitive steps, albeit less thoroughly, whereas strong satisficing refers to situations where the information-processing process is shortened or, in the case of very strong satisficing, not executed at all. This is the case, for example, when the respondent does not even read the question. Respondents who display satisficing behaviour interpret the question in a merely superficial way and report the first answer that comes into their heads or that appears to them to be plausible (Krosnick, Nayaran, & Smith, 1996). This may give rise to response bias, for example when respondents choose the middle response category irrespective of the content of the question in order to simplify the cognitive process.

Three factors determine the likelihood that a respondent will choose a satisficing strategy: (1) task difficulty, (2) cognitive ability, and (3) the motivation of the respondent to fulfil the task (see, among others, Krosnick et al., 1996; Krosnick, 1999b).

The present contribution addresses the most well-known types of response biases in standardised social science surveys – namely, socially desirable responding (Section 2.1), acquiescence (Section 2.2), moderacy response bias and extreme response bias (Section 2.3). The influence of the interview situation is also addressed (Section 3.1: Interviewer effects; Section 3.2: Presence of third parties). So, too, is the design of the questionnaire – that is, the question order (Section 4.1) and the response order (Section 4.2) – which can also cause response bias.

2. Types of response biases

In standardised social science surveys, the most well-known response biases are socially desirable responding, acquiescence, moderacy response bias (i.e., the tendency to use the middle categories on the rating scale), and extreme response bias (i.e., the tendency to use the extreme categories on the rating scale). These biases are addressed in what follows.

2.1 Socially desirable responding

Socially desirable responding (SDR) refers to the tendency of respondents to give predominantly positive self-descriptions (Paulhus, 2002). Expressed in more general terms, it describes the tendency of respondents to comply with what they think the interviewer or other participants expect of them. SDR can take the form of overreporting desirable behaviour or underreporting undesirable behaviour. Respondents are guided in this regard by social norms (Kreuter, Presser, & Tourangeau, 2008). One example is the social norm of voting in elections, which is seen as a civic duty in many Western countries. Failure to vote is deemed to constitute a violation of this norm, which is why – as a comparison with the official turnout figures shows – some respondents falsely report voting (Belli, Traugott, & Beckmann, 2001; Holbrook & Krosnick, 2010).

The tendency to answer in a socially desirable way manifests itself especially in the case of questions about topics that respondents consider to be sensitive (e.g., sexual behaviour, abortion, or drug abuse; Kreuter et al., 2008). The mode of survey data collection also influences the occurrence of SDR. Respondents are more likely to answer in a socially desirable way when the social distance – in the sense of physical proximity or presence – between them and the interviewer is small, as is the case in a face-to-face interview. Here, the respondent runs the risk of being exposed in front of the interviewer. On the other hand, when there is a large social distance between the respondent and the interviewer or researcher, as in the case of a self-administered web survey, for example, the tendency to present oneself in a more positive light is less pronounced (Tourangeau et al., 2000; Tourangeau & Yan, 2007; Kreuter et al., 2008; Holbrook & Krosnick, 2010; Preisendörfer & Wolter, 2014). However, there are also differences between the interviewer-administered modes of survey data collection. Holbrook, Green, and Krosnick (2003) showed that telephone respondents were more likely to exhibit response behaviour that was oriented towards social norms than were face-to-face respondents. The authors attributed this to the fact that a relationship of trust between the interviewer and the respondent could best be built up during a face-to-face interview, which was why respondents were less likely to adapt their answers to social norms in this mode than in telephone interviews (Holbrook et al., 2003). In a validation study, Preisendörfer and Wolter (2014) demonstrated that respondents with a high level of education, females, and older people admitted undesirable behaviour less frequently than did younger respondents, males, and persons with lower education. Krumpal (2013) provides an overview of the determinants of SDR in surveys.

SDR leads to systematic measurement errors that endanger the interpretability of column and row totals, individual differences, and statistical correlations on the basis of these data. For this reason, numerous studies have been devoted to the development of scales for the measurement of social desirability with the aim of employing them to control for the confounding of the data (for an overview of such studies, see, for example, Paulhus, 2002; Kemper, Beierlein, Bensch, Kovaleva, & Rammstedt, 2012). However, many of these scales are too long to be applied as standard in social science surveys. A six-item short scale for the measurement of social desirability (KSE-G), which was developed both in English and in German, offers an economical measurement (Kemper et al., 2012).¹

¹ See also <http://www.gesis.org/kurzskalen-psychologischer-merkmale/kurzskalen/soziale-erwuenschtheit/>.

By using certain “hidden-response” techniques in the case of sensitive questions, it is possible to counteract the occurrence of SDR in the data collection phase. Approaches such as the randomised response technique and the item count technique offer respondents the possibility of giving confidential and anonymous responses. Some studies have shown that, when hidden-response techniques were applied, higher rates of undesirable behaviour were measured than when direct questions were used. However, other studies did not succeed in demonstrating the effectiveness of these techniques. A central problem when applying a hidden-response technique is to explain it to the respondents in such a way that they understand the technique and find it plausible and that they give an honest answer to questions about sensitive topics (for a presentation of the techniques and an overview of the studies available, see, for example, Krumpal, 2013).

2.2 Acquiescence

Acquiescence is one of the response biases that are independent of question content (Krosnick, 1991). Acquiescent respondents agree to statements irrespective of their content. In other words, they select the categories “agree,” “true,” or “yes” (Krosnick, 1999a). This may give rise to systematic distortions (Billiet & Davidov, 2008).

The sources of acquiescence may lie in the stimulus – that is, in the formal design of the question itself – or in the respondent. In the first case, in order to avoid acquiescence, it helps to reword the questions or to ask them in a different way (Billiet & Davidov, 2008). Some studies have found that ambiguous items (McBride & Moran, 1967) or lengthy items (Trott & Jackson, 1967) may be associated with increased acquiescence. It may therefore be helpful to reformulate such items. Ultimately, distortions through acquiescence can be prevented only by avoiding question formats such as “agree-disagree,” “true-false,” and “yes-no” (Krosnick, 1999a).

If one assumes that the cause of acquiescence lies in the respondent, a number of different explanations can be found. First, acquiescence may result from the desire of the respondent to be polite (Leech, 1983, cited in Krosnick, 1999a), as it is more pleasant to agree with something than to disagree. A more far-reaching explanation is that the tendency to acquiesce is a stable personality trait or behavioural tendency (Krosnick, 1991). Krosnick (1991) showed that acquiescence may also result from a satisficing response strategy. In this case, respondents tend to think only of reasons why a statement might be true but not why it might also be false. In other words, they consider a few pro arguments but no contra arguments. This decision-making rule fosters acquiescence, or, to put it another way, people who make an effort to come up with reasons why a statement might not be true or why they might not agree to with it are less susceptible to acquiescence (Krosnick, 1991).

Acquiescence can be identified and controlled with the help of balanced scales (Billiet & Davidov, 2008). In balanced scales, half of the items are worded in the positive direction and the other half are worded in the negative direction. Hence, a statement and its opposite are formulated, thereby creating pairs of mutually exclusive statements. The assumption underlying this approach is that acquiescent respondents will agree to a statement and to its opposite, that these responses will cancel each other out, and that the scale mean will not therefore be biased by acquiescence (Krosnick, 1999a; Billiet & Davidov, 2008).² However, this applies only if positively and negatively formulated items are equally susceptible to acquiescence, which does not appear to be always the case (McClendon, 1992, cited in Billiet & Davidov, 2008). Moreover, the application of balanced scales works only if statements can be

² Balanced scales are applied to reveal acquiescence. However, they cannot prevent acquiescence from distorting the correlations and factor structures of variables (Rammstedt & Farmer, 2013; Rammstedt, Goldberg, & Borg, 2010).

found that are genuinely opposite in meaning and if the respondents display an acquiescent reaction to all items (Krosnick, 1999a).

Data distorted by acquiescence can be corrected using different methods (Krosnick, 1999a). One possibility is to incorporate items into the questionnaire with which acquiescence that lies in the personality of the respondent can be identified. This method is based on the assumption that acquiescent behaviour can be attributed exclusively to the manifestation of a personality trait. However, this is not always so, as has been demonstrated on the basis of the strategic response behaviour satisficing. In such cases, it helps to use balanced scales because, as described above, they enable the scale means to be corrected for acquiescence so that they are not biased.

Who acquiesces? Empirical studies indicate that respondents of lower social status and lower intelligence or cognitive abilities are particularly prone to acquiescence (Krosnick, 1999a). Moreover, it has been demonstrated that acquiescence is higher among respondents with a lower level of formal education (Rammstedt, Goldberg, & Borg; 2010). Like other manifestations of satisficing, the tendency to acquiesce is also influenced by the difficulty of the items and by the respondent's ability and motivation to answer optimally (Krosnick, 1991).

Interviewers also play an important role in the occurrence of acquiescence in telephone and face-to-face interviews. Olsen and Bilgen (2001) found that experienced interviewers obtained more acquiescent reports than did inexperienced interviewers. This may have been due to the fact that experienced interviewers administered the interviews at a faster pace. Moreover, experienced interviewers were more likely to establish a trusting or harmonious relationship with the respondents, which in turn may have elicited greater acquiescence. Accordingly, interviewers should receive training in neutral interview behaviours, in particular (Olson & Bilgen, 2011).

2.3 Moderacy response bias and extreme response bias

Moderacy response bias – that is, the tendency to choose a response category in the middle of the response scale irrespective of question content and of whether the category actually represents the respondent's true attitude – may be an expression of satisficing. Moderacy response bias is reinforced when one middle category is offered, which may result in less thorough responding (Kalton, Robert, & Holt, 1980; Krosnick & Fabrigar, 1997; Schumann & Presser, 1981; Saris & Gallhofer 2007). For an overview of studies on the effects of the middle category in rating scales, see also the *GESIS Survey Guidelines* contribution "Design of Rating Scales in Questionnaires" (Menold & Bogner, 2016).

Extreme response bias is defined as the tendency of a respondent to select the extreme options in rating scales, irrespective of question content (Greenleaf, 1992). It is thus the opposite of moderacy response bias. Although the extreme poles of rating scales are rarely used, this behaviour may occur in some respondents (Hui & Triandis, 1985).

Extreme responding can be measured as the proportion of extreme responses in a series of rating scales. However, there are no uniform findings with regard to two central questions, namely: Is extreme responding a response style or a response set? And is it independent of personal and demographic characteristics (Greenleaf, 1992)?

Greenleaf (1992) came to the conclusion that extreme response style was a stable behaviour over the duration or length of a survey and that it was dependent on the age, the education, and the income of the respondent but not on his or her sex. Older persons, and persons with lower education and lower income, were particularly susceptible to extreme response style.

Hui and Triandis (1985) arrived at the opposite conclusion. They countered the assumption that extreme responding was stable over time with the hypothesis that biases were the result of an interaction

between personality factors and situational characteristics – for example, time-related factors. This hypothesis could be confirmed: The authors found that respondents changed their response behaviour over time or over the duration of a survey. Hence, they concluded that extreme responding was an unstable response behaviour. Hui and Triandis offered several possible explanations for this instability: First, the attempt to answer consistently may ultimately result in the respondent's giving extreme responses at the end of the questionnaire that are not consistent with previous answers. Second, the more familiar the respondent is with the items, and the more certain he or she is about his or her attitude, the more likely it is that extreme responding will occur (Hui & Triandis, 1985).

A further finding is the dependency of extreme responding on the survey mode. Compared to web surveys or postal surveys, telephone interviews and face-to-face interviews have been found to be particularly susceptible to extreme responding but not to extremely negative responses. The finding that telephone and face-to-face respondents are more likely to give extremely positive responses could be linked to the presence of a (telephone) interviewer and to the respondent's aversion to making negative statements in the presence of others (Ye, Fulton, & Tourangeau, 2011).

3. Influence of the interviewer and the interview situation

3.1 Interviewer effects

Like presence effects, interviewer effects are an example of response behaviour that is dependent on the interview situation. The causes lie in the behaviour of the interviewer (Mangione, Fowler, & Louis, 1992) or in characteristics of the interviewer (Haunberger, 2006). Demographic and socio-economic characteristics of the interviewer may influence the respondent's response behaviour, especially when items in the questionnaire address these characteristics (Groves, 1989, cited in Haunberger, 2006). In principle, the presence of an interviewer has positive effects: The interviewer can ask probing questions and is available to answer any questions that the respondent may have; he or she can clarify points that are unclear and can ensure that the respondent understands the question (Mangione et al., 1992). However, interviewers may display behaviours that have negative effects. This is the case, for example, when the interviewer does not read out the response categories in a neutral tone of voice or when he or she does not accurately record the respondent's responses (Mangione et al., 1992).

There are diverse findings regarding the relationship between characteristics of the questionnaire or the questions and the probability that interviewers will influence the response behaviour of the respondents (Mangione et al., 1992). Responses to difficult questions, attitudinal questions, and open-ended questions are considered to be more likely to be influenced by interviewers (Haunberger, 2006). The interviewer's sex and education may influence responses to difficult questions. For example, respondents are more likely to refuse to answer questions about income when they are asked by female interviewers or by interviewers with a higher level of education (Haunberger, 2006).

In the case of attitudinal questions, the findings are contradictory. Some studies have found that attitudinal questions are not more susceptible to interviewer effects than factual questions (Groves & Magilavy, 1986). Other studies have found correlations between the sex of the interviewer and attitudinal questions about topics such as marriage, for example: In contrast to male interviewers, female interviewers received more positive and neutral responses than negative responses (Liu & Stainback, 2013).

Open-ended questions, that is, questions without predefined response categories, favour interviewer effects (Mangione et al., 1992). Haunberger found that the age and education of the interviewer were positively correlated with responses to open-ended questions (Haunberger, 2006). The number of

answers given per open-ended question differs among interviewers. This means that, although the first response to open-ended questions does not necessarily differ among interviewers, there may be differences as to whether the respondent gives a second answer, and this may in turn be due to different behaviours or characteristics of the interviewers (Groves & Magilavy, 1986). Although these effects tend to be minor, they can accumulate within a survey and become statistically significant when the individual interviewer conducts a large number of interviews (Groves & Magilavy, 1986; Mangione et al., 1992).

Interviewer effects may also arise from the interaction between characteristics of the interviewer and the respondent. A study conducted by Schanz (1981) found a positive correlation between attitudinal and behavioural characteristics of interviewers and respondents, for example in relation to child-raising goals or alcohol consumption (Schanz, 1981). Depending on the content of the question, the attitudes and behaviours of an interviewer may therefore have an effect. Respondents may infer attitudes from visible interviewer characteristics, or interviewers may express attitudes non-verbally (Schanz, 1981). On the other hand, performance characteristics of the interviewers, such as response rate and productivity, are not linked to interviewer effects (Groves & Magilavy, 1986).

Various studies show that interviewer training has positive effects on data quality (Haunberger, 2006). As one component of this training, interviewers may be asked to fill in the survey questionnaire themselves. In this way, the interviewers' attitudes to the topic of the survey are revealed, connections between the attitudes of the interviewers and those of the respondents can be recognised and possible influences identified (Haunberger, 2006; Schanz, 1981). At the same time, the interviewers can familiarise themselves with the survey instrument (Haunberger, 2006). A further implication is the fact that it is useful to provide training in probing because the interviewer may influence response behaviour through probing, in particular. Probing occurs most frequently in the case of difficult and open-ended questions. An important starting point is therefore to formulate the questions in such a way that they require as little probing as possible. Pretests can help to improve the formulation of the items (Mangione et al., 1992). For detailed information on interviewer training, see the *GESIS Survey Guidelines* contribution "Interviewer Qualification and Training" (Stiegler & Biedinger, 2016).

3.2 Presence of third parties

Presence effects may occur when third parties are present during the interview and the response behaviour of the respondent is dependent on the presence of these third parties (Lander, 2000). Third parties present may be the spouse or partner, the children, other relatives, or friends or neighbours of the respondent. However, in most cases, it is the spouse or partner (Hartmann, 1994).

Most studies that investigate the effects of the presence of third parties report positive effects on responses to factual questions: Spouses or partners can check whether the reports given are correct (Lander, 2000) and can thus contribute to improving the quality of the data (Zipp & Toth, 2002; Aquilino, 1993). The presence of the spouse or partner may help the respondent to remember (Hartmann, 1994).

However, effects on the validity of the collected data as a result of changed behavioural expectations on the part of all participants – that is, the interviewer, the respondent and third parties present – have also been discussed (Lander, 2000). The presence of third parties (as a methodological problem) has an influence on the data quality because it may have social consequences for the respondent (Hartmann, 1994). Presence effects can therefore be interpreted as the influence of social desirability vis-à-vis the interviewer and third parties (Lander, 2000). When the spouse or partner is present at the interview, differences can be detected in the data. The presence of the spouse or partner may influence responses to relationship-related questions (Lander, 2000). Overall, however, the demonstrable influence of the presence of the spouse or partner is minor, and there is little evidence of social desirability (Zipp &

Toth, 2002). When both spouses/partners are interviewed, the presence of the other spouse or partner leads to increased consensus between the spouses/partners on attitudinal and behaviour-related questions (Zipp & Toth, 2002) and to an increase in traditional responses. However, this increase cannot be attributed solely to the presence of the other spouse or partner. Rather, the presence of the spouse or partner may be a manifestation of a traditional relationship (Lander, 2000). Hence, recent analyses do not show any systematic influences of the presence of third parties, even though inconsistent findings have been reported with regard to individual aspects (Lander, 2000).

Many studies have demonstrated that the presence of a spouse or partner is not coincidental but rather depends on characteristics of both spouses/partners. The characteristics in question explain not only the presence of the spouse or partner but may also influence response behaviour (Hartmann, 1994). The resulting different response behaviour does not necessarily constitute presence effects and biases (Hartmann, 1994). This is precisely because the presence of the spouse or partner and the responses to the questions may be influenced by the same characteristics (Aquilino, 1993). Besides structural reasons (e.g., economic inactivity), the presence of a spouse or partner may also be due to motivational factors on the part of one or both spouses/partners that may also be of relevance in the case of relationship-related questions (Lander, 2000).

These findings have several implications for research practice. It has been found that the longer the interview lasts, the more likely it is that the spouse or partner will be present. Therefore, if one wishes to interview the respondent alone, it is better to keep the interview short (Hartmann, 1994). Interviewers should be instructed to collect information about the presence of third parties (Aquilino, 1993). While the fact that third parties were present is often known, the reasons and motives for their presence are not (Hartmann, 1994). Overall, it can be stated that the presence of the spouse or partner results in hardly any response biases and may help the respondent to remember. Nonetheless, researchers should be aware of the context of the interview when the interaction between the interviewer and the respondent is extended by third parties who are present during the interview (Zipp & Toth, 2002).

4. Influence of the questionnaire

4.1 Question-order effects

Not only the question wording but also the order in which the questions are asked may influence the way in which the respondent interprets and evaluates them. This is because previous questions constitute the context for subsequent questions (see, among others, Tourangeau et al., 2000; Moore, 2002; Krosnick & Presser, 2010). When answering subsequent questions, respondents may orient themselves towards previous questions or towards their responses to these questions (so-called halo effects), which gives rise to a systematic distortion of the responses (Schnell, Hill, & Esser, 2011; Schnell, 2014).

Previous questions may trigger the retrieval of certain information, which is then very easily accessible to the respondent when answering subsequent questions about similar topics. As a result, respondents probably build their responses to subsequent questions on the same deliberations as in previous questions about similar topics. For example, if a respondent is asked how satisfied he is with his marriage, this increases the probability that he will also take his marriage into account when answering the question about his satisfaction with life in general. This is known as an assimilation effect. It is the opposite of a contrast effect, where responses to subsequent questions are steered in the opposite direction to the previous questions: In order to avoid repetitions when answering subsequent questions,

respondents may deliberately exclude information from the response process that they used in the case of previous questions (Schwarz & Bless, 1992; Mason, Carlson, & Tourangeau, 1994).

To identify the occurrence of question-order effects, the order of the questions can be varied in the pretest or in a split ballot design. However, the effort needed to construct the questionnaire increases as a result.

4.2 Response order

The response order of a list or a rating scale may influence the response that the respondent chooses. In the case of the visual presentation of the response alternatives – for example, on showcards or in self-administered surveys such as web surveys – respondents tend to choose one of the first alternatives (primacy effect). The situation is different in the case of response alternatives that are orally presented in telephone or face-to-face interviews. Here, respondents tend to choose from the last-mentioned alternatives (recency effect; Krosnick & Alwin, 1987; Krosnick, 1991; Schwarz, Hippler, & Noelle-Neumann, 1992; Bishop & Smith, 2001).

Two theoretical models have proved their worth in explaining the occurrence of response-order effects: the cognitive elaboration model of response-order effects (Sudman, Bradburn, & Schwarz, 1996) and satisficing theory (Krosnick et al., 1996). Following the cognitive elaboration model, the occurrence of response-order effects basically depends on whether the respondent has a chance to think about the content of the individual response alternatives. This chance is determined by three factors: (1) the serial position in which an alternative is presented (at the beginning, in the middle, or at the end); (2) the presentation mode (orally, by reading it out, or visually, in self-administered surveys or via showcards during interviews), and (3) the plausibility of a response alternative. According to the assumptions of the model, primacy effects are expected in the case of visually presented response formats because respondents have a better chance to think about the first alternatives. This does not apply when the first alternatives appear implausible to the respondent. In this case, recency effects are expected for visual formats, too. Recency effects are expected in the case of orally presented response formats because respondents hardly have a chance to cognitively process the first alternatives that are read out because this process is interrupted by the naming of the next alternatives. Here, too, it is assumed that, if alternatives appear implausible to the respondent, primacy effects occur also in the case of oral presentation (Bishop & Smith 2001).

From the perspective of satisficing theory, response-order effects occur when, for reasons of low motivation, low cognitive abilities, or high task difficulty, respondents wish to simplify the cognitive response process. In this case, they use the first acceptable response alternative without paying particular attention to the other alternatives (Krosnick & Alwin, 1987; Krosnick, 1991). The expected direction of the effects is then the same as in the cognitive elaboration model of response-order effects – that is, primacy effects in the case of visually presented formats and recency effects in the case of orally presented formats (Bishop & Smith, 2001).

Galesic, Tourangeau, Couper, and Conrad (2008) used the eye-tracking method to analyse respondents' eye movements and fixation times when answering a web questionnaire. The authors identified two response behaviours of respondents that can be identified as an expression of strong and weak satisficing, respectively. Some respondents chose alternatives from the first half of the list without even fixing on the lower alternatives (strong satisficing). Other respondents showed significantly longer fixation times for the upper half of the list than for the lower half irrespective of the content of the response alternatives, which is an indication of more thorough cognitive processing of the upper alternatives (weak satisficing). Moreover, the current state of research shows that primacy effects are more likely to occur when the list of response alternatives is longer (Galesic et al., 2008), the cognitive

abilities of respondents are lower (Krosnick & Alwin, 1987), and the respondents are older (Knäuper, 1999). However, education does not influence response-order effects (Knäuper, 1999).

On the basis of data from 548 experiments that were conducted during telephone surveys, Holbrook, Krosnick, Moore, and Tourangeau (2007) examined the characteristics of questions and respondents that favoured the occurrence of recency effects. The authors found that recency effects occurred most in the case of complicated questions. This effect was more pronounced among respondents with the least education when the questions were more complicated, when they were located at the end of the questionnaire, and when the response alternatives were formulated as complete sentences rather than as words or phrases. When questions implicitly or explicitly encouraged respondents to wait until they had heard all the response alternatives before formulating a response, recency effects were nevertheless more common (Holbrook et al., 2007).

In order to counter response-order effects, Holbrook et al. recommended that, when developing a question instrument, long response lists should be avoided where possible. Otherwise, the following measures can be considered: In the case of lists with unordered response alternatives, these alternatives can be routinely rotated so that different alternatives appear at the top or the bottom, thereby avoiding bias in the data (Holbrook et al., 2007). While this procedure is easily implemented in computer-assisted surveys, its implementation in paper surveys is very burdensome and costly. Nor does it solve the problem that – rather than giving an answer that reflects their true attitude – some respondents still give answers that have been influenced by response-order effects. Therefore, a further possible measure is to increase the motivation of the respondent in order to prevent satisficing and counter the occurrence of response-order effects. This can be done by means of instructions, for example, informing respondents that the question they are about to answer is difficult and requires full concentration (Krosnick & Alwin, 1987). The general recommendation that questions should be formulated as simply as possible and that unusual or abstract terms should be avoided applies also with regard to response-order effects. The rationale is that questions that are kept simple also keep the cognitive burden on respondents to a minimum, which in turn counters the occurrence of satisficing (Holbrook et al., 2007). Moreover, forced-choice response formats are preferable to check-all response formats because fewer primacy effects and deeper information processing on the part of respondents have been observed when forced-choice formats were used (Smyth, Dillman, Christian, & Stern, 2006). And finally, one could consider applying rating scales for the evaluation of each individual response alternative rather than using a list of alternatives. However, one would have to weigh up whether this might increase the length of the questionnaire to such an extent that other negative effects might occur. Further recommendations for the design of response formats can be found in the *GESIS Survey Guidelines* contribution “Question Formulation” (Lenzner & Menold, 2016).

References

- Aquilino, W. S. (1993). Effects of spouse presence during the interview on survey responses concerning marriage. *Public Opinion Quarterly*, 57, 358-376.
- Belli, R. F., Traugott, M. W., & Beckmann, M. N. (2001). What leads to voting overreports? Contrasts of overreporters to validated votes and admitted nonvoters in the American Election Studies. *Journal of Official Statistics*, 17, 479-498.
- Billiet, J. B., & Davidov, E. (2008). Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design. *Sociological Methods & Research*, 36, 542-562.
- Bishop, G., & Smith, A. (2001). Response-order effects and the early Gallup split-ballots. *Public Opinion Quarterly*, 65, 479-505.

- Burden, B. C. (2008). The social roots of the partisan gender gap. *Public Opinion Quarterly*, 72(1), 55-75.
- Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. In: S. Leinhardt (Ed.), *Social Methodology*. San Francisco: Jossey-Bass Publishers.
- Galesic, M., Tourangeau, R., Couper, M., & Conrad, F. (2008). Eye tracking data. New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, 72(5), 892-913.
- Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly*, 56, 328-351.
- Groves, R. M., & Magilavy, L. J. (1986). Measuring and explaining interviewer effects in centralized telephone surveys. *Public Opinion Quarterly*, 50(2), 251-266.
- Groves, R. M., Fowler, J. F., Couper, M., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. Hoboken: John Wiley & Sons.
- Hartmann, P. (1994). Interviewing when the spouse is present. *International Journal of Public Opinion Research*, 6(3), 298-306.
- Haunberger, S. (2006). Das standardisierte Interview als soziale Interaktion: Interviewereffekte in der Umfrageforschung. *ZA-Information*, 58, 23-46.
- Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports. Tests using the item count technique. *Public Opinion Quarterly*, 74(1), 37-67.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires. *Public Opinion Quarterly*, 67, 79-125.
- Holbrook, A. L., Krosnick, J. A., Moore, D., & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally. The impact of question and respondent attributes. *Public Opinion Quarterly*, 71(3), 325-348.
- Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly*, 49, 253-260.
- Kahn, R. L., & Cannell, C. F. (1957). *The Dynamics of Interviewing: Theory, Technique, and Cases*. New York: Wiley.
- Kalton, G., Robert, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. *Statistician*, 29, 65-78.
- Kemper, C. J., Beierlein, C., Bensch, D., Kovaleva, A., & Rammstedt, B. (2012). Eine Kurzsкала zur Erfassung des Gamma-Faktors sozial erwünschten Antwortverhaltens. Die Kurzsкала Soziale Erwünschtheit-Gamma (KSE-G). *GESIS-Working Papers*, No. 2012/25. Cologne: GESIS Leibniz Institute for the Social Sciences.
- Knäuper, B. (1999). The impact of age and education on response order effects in attitude measurement. *Public Opinion Quarterly*, (63), 347-370.
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys. *Public Opinion Quarterly*, 72(5), 847-865.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5, 213-236.
- Krosnick, J. A. (1999a). Maximizing questionnaire quality. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Political Attitudes* (pp. 37-57). San Diego: Academic Press.
- Krosnick, J. A. (1999b). Survey research. *Annual Review of Psychology*, 50, 537-567.

- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, *51*(2), 201-219.
- Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey Measurement and Process Quality* (pp. 141-164). New York: Wiley-Interscience.
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden, & J. D. Wright (Eds.), *Handbook of Survey Research* (pp. 263-313). Bingley: Emerald Group Publishing.
- Krosnick, J. A., Narayan, S. S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. In M. T. Braverman, & J. K. Slater (Eds.), *Advances in Survey Research* (pp. 29-44). San Francisco: Jossey-Bass.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality and Quantity*, *47*, 2025-2047.
- Lander, B. (2000): Anwesenheitseffekte im Wandel. Eine Sekundäranalyse zur Anwesenheit des Partners im Interview anhand des ALLBUS 1980 bis 1998. *Zeitschrift für Soziologie*, *29*(3), 227-239.
- Lenzner, T., & Menold, N. (2016). Question Formulation. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_017
- Liu, M., & Stainback, K. (2013). Interviewer gender effects on survey responses to marriage related questions. *Public Opinion Quarterly*, *77*(2), 606-618.
- Mangione, T. W., Fowler, F. J., & Louis, T. A. (1992). Question characteristics and interviewer effects. *Journal of Official Statistics*, *8*(3), 293-307.
- Mason, R., Carlson, J. E., & Tourangeau, R. (1994). Contrast effects and subtraction in part-whole questions. *Public Opinion Quarterly*, *58*, 569-578.
- McBride, L., & Moran, G. (1967). Double agreement as a function of item ambiguity and susceptibility to demand implications of the psychological situation. *Journal of Personality and Social Psychology*, *6*, 115-118.
- 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
- Moore, D. W. (2002). Measuring new types of question-order effects. *Public Opinion Quarterly*, *66*(1), 80-91.
- Olson, K., & Bilgen, I. (2011). The role of interviewer experience on acquiescence. *Public Opinion Quarterly*, *75*(1), 99-114.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah NY: Erlbaum.
- Preisendörfer, P., & Wolter, F. (2014). Who is telling the truth? A validation study on determinants of response behavior in surveys. *Public Opinion Quarterly*, *78*(1), 126-146.
- Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*, 1137-1145.
- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, *44*, 53-61.
- Rasinski, K. A. (1989). The effect of question wording on public support for government spending. *Public Opinion Quarterly*, *53*, 388-394.

- Schanz, V. (1981). Interviewereffekte: zusammenfassende Darstellung von Arbeiten, die im Rahmen zweier von ZUMA betreuter Projekte entstanden sind. *ZUMA Nachrichten*, 5(9), 36-46.
- Schnell, R. (2014). *Survey-Interviews. Methoden standardisierter Befragungen*. Wiesbaden, VS Verlag für Sozialwissenschaften.
- Schnell, R., Hill, P. B., & Esser, E. (2011). *Methoden der empirischen Sozialforschung*. Munich, Oldenbourg Verlag.
- Schwarz, N., & Bless, H. (1992). Assimilation and contrast effects in attitude measurement: An inclusion/exclusion model. *Advances of Consumer Research*, 19, 72-77.
- Schwarz, N., Hippler, H.-J., & Noelle-Neumann, E. (1992). A cognitive model of response order effects in survey measurement. In N. Schwarz, & S. Sudman (Eds.), *Context Effects in Social and Psychological Research*. New York: 187-201.
- Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Comparing check-all and forced-choice question formats in web surveys *Public Opinion Quarterly*, 70, 66-77.
- Stiegler, A., & Biedinger, N. (2016). Interviewer Skills and Training. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_013
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859-883.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Trott, D. M., & Jackson, D. N. (1967). An experimental analysis of acquiescence. *Journal of Experimental Research in Personality*, 2, 278-288.
- Ye, C., Fulton, J., & Tourangeau, R. (2011). Research synthesis. More positive or more extreme? A meta-analysis of mode differences in response choice. *Public Opinion Quarterly*, 75(2), 349-365.
- Zipp, J. F., & Toth, J. (2002). She said, he said, they said. The impact of spousal presence in survey research. *Public Opinion Quarterly*, 66, 177-208.