

The Effect of Question Positioning on Data Quality in Web Surveys

Neuert, Cornelia

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Neuert, C. (2024). The Effect of Question Positioning on Data Quality in Web Surveys. *Sociological Methods & Research*, 53(1), 279-295. <https://doi.org/10.1177/0049124120986207>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft

The Effect of Question Positioning on Data Quality in Web Surveys

Sociological Methods & Research
2024, Vol. 53(1) 279–295
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0049124120986207
journals.sagepub.com/home/smr



Cornelia Eva Neuert¹ 

Abstract

The quality of data in surveys is affected by response burden and questionnaire length. With an increasing number of questions, respondents can become bored, tired, and annoyed and may take shortcuts to reduce the effort needed to complete the survey. In this article, direct evidence is presented on how the position of items within a web questionnaire influences respondents' focus of attention. In two experiments, part of an eye-tracking study and an online survey, respectively, a variety of indicators show that data quality is lower if the experimental question is positioned at the end rather than at the beginning of a questionnaire. Practical implications are discussed.

Keywords

data quality, question positioning, response bias

Introduction and Background

In survey research, it is of utmost significance that survey responses are accurate and of high data quality, that is, valid, reliable, and unbiased (Fowler 2013; Fowler and Cannell 1996; Schober and Conrad 1997).

¹ GESIS — Leibniz Institute for the Social Sciences Mannheim, Germany, Mannheim, Germany

Corresponding Author:

Cornelia Eva Neuert, GESIS — Leibniz Institute for the Social Sciences Mannheim, Germany, Mannheim, Germany.

Email: cornelia.neuert@gesis.org

According to the question–answer process, respondents are required to go through the following complex cognitive processes while answering (Tourangeau, Rips, and Rasinski 2000): Respondents must comprehend the question, retrieve information, make a judgment, and give a response. High-quality data can only be expected when respondents engage in all four steps of question answering (termed optimizing). Satisficing, in contrast, is characterized by respondent behavior that does not intend to address one or even more of the four stages thoroughly, but to take shortcuts to reduce the effort needed; for instance, by not differentiating among response options, selecting non-substantive answers like “don’t know,” or rushing through surveys (Krosnick 1991; Krosnick and Alwin 1987; Zhang and Conrad 2014).

The likelihood of satisficing depends on task difficulty (e.g., question difficulty), respondents’ cognitive ability, and respondents’ motivation (Biemer and Lyberg 2003; Cannell, Miller, and Oksenberg 1981; Groves et al. 2009; Krosnick 1991; Schwarz and Hippler, 1991). Survey satisficing is thus fostered by survey (design) characteristics that are likely to increase the burden of responding, hence reducing data quality. Or to put it differently: If the respondent burden can be conducive to motivating optimization when answering questions, then eliminating or lowering that burden might mitigate satisficing, and thus increase data quality. One such survey design property is survey length and, related to the overall length of a survey, the position of a question in the questionnaire.

In previous research, respondent burden has often been measured in terms of the time it takes to complete the questionnaire. It has been shown that respondent burden increases with questionnaire length, while respondent motivation decreases (Groves et al. 1999; Singer et al. 1999). The perceived burden indicated by the announced survey length, for example, influenced whether respondents were willing to take part in a survey (Galesic and Bosnjak 2009; Hansen 2007). Respondents were also more likely to answer “no” to questions that allowed them to skip following questions and thus to save time (Shields and To 2005). Apart from effects on unit and item nonresponse, the effect of survey length on data quality in later parts of the survey is also a worrying factor to be considered. Items placed later in a questionnaire may additionally be affected by respondents’ increased fatigue: Their experienced burden might increase with every preceding question (Galesic 2006). Thus, questions that are positioned at the end of a questionnaire may be processed differently than questions that are asked at the beginning of a survey (Helgeson and Ursic 1994). Previous research regarding question

placement effects—whether earlier or later in a questionnaire—on data quality is mixed, and mostly based on indirect indicators of measurement error (Peytchev and Peytcheva 2017). Clancy and Wachsler (1971) did not find increased tendencies of “yes-saying” (which was used as proxy for boredom) or willingness to report socially undesirable behaviors when questions were asked at the end of a 25-minute questionnaire, rather than at the beginning, and no other effects which could conceivably be ascribed to boredom (such as “don’t know” answers). Sudman and Bradburn (1974) reported that the position of a question had little (for behavioral questions) or negligible effects (for attitudinal questions). Johnson, Sieveking, and Clanton (1974) found that college students gave fewer responses to open-ended questions when they were asked towards the end of the survey. In contrast, the results of a study of Burchell and Marsh (1992) pointed in the opposite direction, with more responses at the end of a lengthy questionnaire. Several other studies have found that data quality was lower for questions positioned later in a questionnaire compared to the same items placed earlier. These studies reported higher rates of missing data and “don’t know” answers, greater tendencies to respond with agreement, shorter answers to open-ended questions, shorter response times, or less differentiation among items (also called “straightlining”) when they appeared later in a questionnaire (Galesic and Bosnjak 2009; Herzog and Bachman 1981; Krosnick et al. 2002), as well as lower criterion validity (Peytchev and Peytcheva 2017). However, in most of these studies, findings are based on indirect indicators for respondent burden and data quality.

In this study, eye-tracking methodology was used as a more direct indicator of respondent burden to analyze differences in the depth of information processing (Galesic et al. 2008; Galesic and Yan 2011). Eye-tracking allows researchers to observe respondents’ eye fixation patterns: which words are fixated, how long are they fixated, and the sequence of fixations (Romano Bergstrom and Schall 2014). These measures can be used to understand the underlying comprehension processes (Just and Carpenter 2018) and to examine whether there are differences in the depth of information processing and attention depending on the position of a question within a survey.

Research Questions

The following research questions are addressed:

1. What effects do variations in question positioning within the questionnaire have on data quality?
2. Does the positioning influence depth of processing? Will a question positioned at the end of a questionnaire receive the same amount of attention as it will at the beginning?

The following hypotheses are postulated:

Hypothesis 1: Data quality will be lower when the question is asked later rather than earlier in the questionnaire.

Hypothesis 2: Later positioning will negatively affect respondents' motivation to invest sufficient attention and effort in answering.

Specifically, I expect that item response rates, response times, and differentiation of responses to items will be lower if the question is positioned at the end of the questionnaire, when compared to asking it at the beginning. Furthermore, I expect that the eye-tracking measures of cognitive effort and attention will be lower if the question is positioned at the end of the questionnaire compared to asking it at the beginning, when comparing how much time respondents spend looking at the question.

Experiment I

Data and Experimental Design

The first experiment was embedded in a laboratory eye-tracking study ($N = 131$) which was conducted at GESIS—Leibniz Institute for the Social Sciences in Mannheim (Germany) in April and May of 2017. In total, 131 respondents (50 percent female, 65 percent between 18 and 34 years old, 22 percent between 35 and 54 years old, 13 percent 55 years or older, 29 percent with 10 years of schooling or less, and 71 percent with at least 12 years' schooling) participated in the study. Respondents were recruited from the respondent pool maintained by the institute or by word of mouth with quotas for sex. No quotas for age and education were set, but it was intended to obtain a sample as diverse as possible. Respondents answered the web survey while their eye movements were tracked with a Senso Motoric Instruments (SMI) RED250 mobile Eye Tracker, which is a mobile device that was mounted on the bottom frame of a 17" TFT monitor (resolution 1280x1024). At the beginning of the questionnaire, all respondents answered two warm-up questions, which were used to calculate their fixation and reading rate. These measures were later used as covariates in the statistical

analyses of fixation count and fixation time to control for interindividual differences. The questionnaire included several unrelated experiments with questions on a variety of topics such as politics, social inequality, media use, health, and personality traits. All respondents received a compensation of €20.

To examine the effect of question positioning on the quality of responses, I used a scale on impulsive behavior (Kovaleva et al. 2014; Whiteside and Lynam 2001) consisting of eight rating scale items. Two items each measure one of the four different constructs of impulsive behavior: urgency, (lack of) premeditation, (lack of) perseverance, and sensation seeking. The question was presented in a grid question format. Respondents answered the question using a fully -labeled five-point unipolar agreement scale ranging from “doesn’t apply at all” to “applies fully.” The question wording (in English and German) can be found in Online Appendix A (which can be found at <http://smr.sagepub.com/supplemental/>).

A between-subject design was used with random assignment of respondents to two experimental conditions. Half of the respondents answered the experimental question at the beginning of the survey, directly after the two warm-up questions. The other half answered the question at the end of the survey before the sociodemographic variables. Up to that point, the average completion time was approximately 30 minutes. During web survey completion, respondents’ eye movements were tracked. Bivariate analyses on sex ($\chi^2 = 1.72$; $df = 1$; $p = .19$), age ($\chi^2 = .097$; $df = 2$; $p = .95$), and education ($\chi^2 = 2.38$; $df = 2$; $p = .30$) indicated that the randomization of respondents to the experimental conditions had been successful.

Method

To test the effect of question positioning on the quality of responses, differences in means and proportions for the following indicators were compared.

Item nonresponse. Item nonresponse is defined as missing information (gaps) in a data set, that is, the proportion of respondents giving no response to an item.

Response differentiation (Straightining). Response differentiation reflects the extent to which respondents differentiate in their answers to the eight items

(response variation). Nondifferentiation occurs when respondents do not differentiate in their responses but give (nearly) identical responses to all items. The indicator probability of differentiation P_d is calculated according to Krosnick and Alwin (1988) as $P_d = 1 - \sum_{i=1}^n P_i^2$, where P_i is the proportion of the values rated on a given point of a response scale and n is the number of rating points. If $P_d = 0$ this means that all items were answered by selecting the same response, while a higher P_d means that different response options were used, hence more differentiation.

Response time. Response times are measured in milliseconds using a client-side paradata script (UCSP, Kaczmirek 2005).

Cognitive effort and attention. To analyze potential differences in the cognitive effort and attention invested by respondents, the eye-tracking metrics fixation time and fixation count are used as indicators. Eye tracking is used to record the respondents' eye movements while they answer an online questionnaire on a computer. Eye-tracking enables us to examine respondents' attention and cognitive processing for specific parts of the question. The connection between eye movements and cognitive processing is based on the *immediacy assumption* and the *eye-mind assumption* (Just & Carpenter, 1980). These two assumptions state that eye movements provide direct information about what respondents are currently processing and how much cognitive effort is involved. To analyze their eye movements, so called "areas of interest" (AOIs) are defined, for which the eye-tracking metrics fixation time and fixation count can be calculated. Fixation time is the total duration of fixations on a respective AOI. Fixation count is the total number of fixations on a specific AOI. I defined six different AOIs: one covering the whole question, one covering the question text, one covering the input fields, one covering the item texts of all eight items, and two covering the first half and the second half of the item list, respectively (see Figure 1 in Online Appendix B [which can be found at <http://smr.sagepub.com/supplemental/>]). Following the assumptions outlined above, the time spent fixating on an AOI reflects the time on processing that area (Just and Carpenter, 1980). Consequently, longer fixation duration can be interpreted as deeper processing of that area of the question (Craik 1999; Joos, Rötting, and Velichkovsky 2003). Due to technical difficulties or poor recording quality, the eye-tracking data of eleven participants (8.4 percent) had to be excluded from the analysis.

Response means. For each of the eight items, I additionally tested whether the means between the early positioning and the late positioning condition differed significantly.

All statistical analyses were performed with SPSS 24. To compare response differentiation, response times, and response means, analyses of variance (ANOVAs) were conducted. For response means, a Bonferroni adjustment for multiple comparisons was applied. To analyze the eye-tracking metrics fixation time and fixation count, analyses of covariance (ANCOVAs) were conducted with reading rate and fixation rate as covariates to control for interindividual differences.

Results

With regard to *item nonresponse*, the analyses showed that all respondents answered all eight items and no item nonresponse occurred at all. Comparing the degree of *response differentiation*, respondents in the early positioning condition differentiated significantly more in their answers ($M_{\text{Early}} = .641$) compared to respondents receiving the items at the end of the questionnaire ($M_{\text{Late}} = .599$; $F(1, 130) = 4.451$; $p = .037$). Respondents also took markedly more time to answer the items when they were asked early in the questionnaire ($M_{\text{Early}} = 60.88$) than when they were presented at the end of the questionnaire ($M_{\text{Late}} = 45.99$; $F(1, 130) = 16.815$; $p < .001$). Finally, I investigated how much *cognitive effort and attention* respondents were willing to invest when answering the scale on impulsive behavior. In line with the second hypothesis, Table 1 shows that respondents took more time to answer the complete question when it was presented at the beginning of the questionnaire than when it was positioned late in the questionnaire ($F(1, 119) = 9.429$, $p = .003$). When the question was further distinguished into the areas of the question text and the items, Table 1 shows that respondents took more time to read the question text ($F(1, 119) = 18.551$, $p < .001$) and to answer the question on the input field ($F(1, 119) = 5.881$, $p = .017$) when it was presented early rather than late. Fixation times on the area of the eight items were not significantly different across conditions ($F(1, 120) = 3.059$, $p = .083$). Considering fixation count, respondents fixated more on the whole question ($F(1, 119) = 7.909$, $p = .006$), on the question text ($F(1, 119) = 17.417$, $p < .001$), and on the list of items ($F(1, 119) = 4.149$, $p = .044$), when it was presented early rather than late. To further investigate whether item placement has an effect on the processing time, I distinguished the list of items into two halves, and compared the fixation time spent on the first and the second

Table 1. Means and Standard Errors (in parentheses) of Fixation Time and Count for the Scale of Impulsive Behavior across Conditions.

Eye-Tracking Parameter	Question Part	Early Position	Late Position
Fixation time (in seconds)	Question total	47.04 (1.70)	39.33 (1.81)
	Question text	4.15 (0.27)	2.44 (0.29)
	Items total	20.18 (0.90)	17.86 (0.96)
	Items first half	11.51 (0.57)	10.64 (0.61)
	Items second half	8.47 (0.44)	7.01 (0.47)
	Scale total	20.79 (0.89)	17.60 (0.95)
Fixation count	Question total	169.14 (5.91)	144.53 (6.32)
	Question text	18.41 (1.08)	11.75 (1.15)
	Items total	86.10 (3.38)	75.88 (3.62)
	Items first half	50.30 (2.21)	45.76 (2.37)
	Items second half	34.93 (1.59)	29.04 (1.71)
	Scale total	57.99 (2.62)	51.55 (2.80)

Note: To control for interindividual differences, respondents reading rate and fixation rate was used as a covariate in the statistical analyses. The table displays estimated marginal means after controlling for the covariates, respectively.

half of the list of items. The findings showed that the differences in fixation time and count were not significant for the first four items [Time: $F(1, 119) = 1.107, p = .295$]; Count: ($F(1, 119) = 1.907, p = .170$)], but for the second half of items [(Time: $F(1, 119) = 4.967, p = .028$); Count: ($F(1, 119) = 6.205, p = .014$)].

Table 2 presents the response means of the eight items. After applying the Bonferroni correction, no significant differences in means between the early versus late positioning condition were observed.

Table 2. Means and Standard Deviations for the Eight Items of the Scale of Impulsive Behavior across Conditions.

	Early Position			Late Position		Difference	F	p
	Mean	SD		Mean	SD			
Item 1	2.73	.99	<	2.95	.98	-.23	1.750	.188
Item 2	2.39	1.04	<	2.75	.97	-.36	4.215	.042
Item 3	3.76	.90	>	3.49	.90	.27	2.848	.094
Item 4	3.74	.83	>	3.38	.96	.36	5.200	.024
Item 5	4.03	.72	>	3.85	.89	.18	1.697	.195
Item 6	3.35	1.09	>	3.23	1.06	.12	.394	.531
Item 7	3.33	1.01	<	3.37	.93	-.04	.045	.833
Item 8	3.56	.91	>	3.43	.98	.13	.613	.435
n	66			65				

Note: One-way analyses of variance (ANOVAs) for differences in means ($p < .05$). Tests were adjusted for multiple comparisons between conditions using the Bonferroni α -correction method. With this correction, for the eight comparisons made here, a p -value of $<.00625$ is needed to reach significance. Test statistics that were statistically significant before this correction was made are in italics.

Experiment 2

To verify the robustness of the results with a larger sample and to gain some further in-depth insights whether the effect of the position on response behavior could be replicated in a shorter survey that is within the recommended survey length for web surveys (Revilla and Ochoa 2017), the experiment was repeated within a shorter survey.

Data and Experimental Design

The second study was conducted with respondents from the German opt-in online panel of the Respondi AG. The experiment was included in a web survey on society and politics, which was fielded between November 15 and November 20, 2018. The web survey used quotas for sex, age, and education. Respondents received a small incentive in form of points in exchange for participating in the survey. Overall, 1,071 panelists accepted the survey invitation, of whom 111 were screened out, 91 broke off, and 869 completed the survey. The break-off rate was 9 percent (cf. Callegaro and DiSogra 2008). The mean duration of the questionnaire completion was 8 minutes 27 seconds (median: 6 minutes 55 seconds).

The same scale as in experiment 1 was used as the experimental question, that is, the scale on impulsive behavior comprising eight items. Again, a between-subject design with random assignment to two experimental conditions was used. Half of the respondents answered the experimental question in the beginning of the survey, after three socio demographic questions and two warm-up questions (after 4 pages, 5 items). The other half answered the question at the end of the survey before an evaluation question (after 27 pages, 42 items). The question was presented as a grid question.

To evaluate possible differences in the sample composition between the two treatment groups, χ^2 -tests were conducted. The results showed no statistically significant differences for age ($\chi^2 = 1.70$; $df = 5$; $p = .889$) and education ($\chi^2 = 2.17$; $df = 2$; $p = .338$). However, the distribution of sex was not equal across conditions ($\chi^2 = 7.54$; $df = 1$; $p = .006$), with more women than men in the “early” compared to the “late” condition. The variable sex was included in the analyses as a covariate.

Method

The second web-based experiment used the same indicators of data quality as experiment 1, with two exceptions. First, the eye-tracking metrics fixation time and fixation count could not be used as indicators for measuring cognitive effort and attention. Second, response times were measured using server-side response times (in seconds). Since it is not possible in an online survey to ensure that respondents complete the survey without interruption, I have excluded response time outliers from the analyses. I excluded those respondents who had response times below or above the median plus/minus the upper and lower quartile range multiplied by three [Hoaglin, Mostellar and Turkey 2000; lower threshold: $Q_{.50} - (3x(Q_{.50} - Q_{.25}))$; upper threshold: $Q_{.50} + (3x(Q_{.75} - Q_{.50}))$].¹

All statistical analyses were performed with SPSS 24. Pearson's X^2 -tests of independence were conducted for the comparison of non-response across conditions. To compare response differentiation and response means, one-way ANCOVAs were conducted with sex included as covariate. As in experiment 1, Bonferroni correction was applied to adjust for multiple comparisons when comparing response means.

Results

In contrast to the eye-tracking study, respondents did not provide an answer to all items of the experimental question in the web survey. However, the level

Table 3. Means and Standard Errors for the Eight Items of the Scale on Impulsive Behavior across Conditions.

	Early Position			Late Position		Difference	F	p
	Mean	SD		Mean	SD			
Item 1	2.54	.94	<	2.81	.97	-.27	18.868	.000
Item 2	2.22	.98	<	2.53	1.04	-.31	20.161	.000
Item 3	3.73	.87	>	3.59	.87	.14	5.686	.017
Item 4	3.74	.87	>	3.59	.88	.15	6.788	.009
Item 5	4.07	.79	>	3.87	.84	.20	12.042	.001
Item 6	3.74	.96	>	3.62	.96	.12	2.987	.084
Item 7	2.75	1.01	<	3.00	.99	-.24	9.305	.002
Item 8	2.91	.97	<	3.10	.95	-.18	6.099	.014
n	439			430				

Note: One-way analyses of covariance (ANCOVAs) for differences in means ($p < 0.05$). Tests were adjusted for multiple comparisons between conditions using the Bonferroni α -correction method. With this correction, for the eight comparisons made here, a p -value of $<.00625$ is needed to reach significance. Test statistics that were statistically significant before this correction was made are in italics, statistics in boldface remained significant after the adjustment for multiple comparisons.

of item nonresponse was still relatively low. For items 2, 3, 4, and 5, item nonresponse rates varied between 0.5 percent and 0.9 percent in the early condition and between 0.2 percent and 0.7 percent in the late condition. All respondents in the early condition gave a response to item 7, while 0.7 percent of the respondents in the late positioning condition did not answer this item. Across conditions, we found no item nonresponse in item 1, item 6, and item 8. Overall, the item nonresponse rates did not differ significantly depending on an early or late position in the questionnaire (see Table A.1 in Online Appendix C [which can be found at <http://smr.sagepub.com/supplemental/>]).

When considering the level of *response differentiation*, the results showed that respondents in the early positioning condition differentiate their answers to a significantly higher degree ($M_{\text{Early}} = .597$) compared to respondents receiving the items at the end of the questionnaire ($M_{\text{Late}} = .549$; $F(1, 868) = 17.355$; $p < .001$). Respondents spent significantly more time answering the questions when they were presented in the beginning ($M_{\text{Early}} = 46.6$) compared to at the end of the questionnaire ($M_{\text{Late}} = 39.4$; $F(1, 804) = 13.76$; $p < .001$).

Table 3 shows the response means of the items on impulsive behavior. After applying the Bonferroni correction, there were significant differences in means between the early versus late positioning condition for item 1

(“*Sometimes I do things impulsively that I shouldn’t do*”; $F(1, 868) = 18.868, p < .001$), item 2 (“*I sometimes do things to cheer myself up that I later regret*”; $F(1, 861) = 20.161, p < .001$), item 5 (“*I always bring to an end what I have started*”; $F(1, 865) = 12.042, p = .001$), and item 7 (“*I am willing to take risks*”; $F(1, 865) = 9.305, p = .002$). The response distributions were similar to those in the first experiment.

Discussion and Conclusions

In survey research, researchers face a trade-off between high survey costs and trying to get as much as possible out of a survey, which can come at the expense of the quality of the data collected. With increasing survey length, the risk of respondents becoming tired and burdened grows, as does the likelihood of them reducing their effort by applying strategies of satisficing response behavior. In the first experiment, I was able to investigate the depth of processing of survey questions depending on their position in the questionnaire by using eye tracking. The results showed differences in data quality when questions are asked at the end, rather than at the beginning of the questionnaire, indicating satisficing response behavior. These differences can be explained by the amount of time people spend processing: The experimental question received less fixation time, and hence less attention, when it was presented at the end of the questionnaire than when it was presented at the beginning.

The first experiment had some limitations. First, it was a laboratory study with a relatively small sample size—albeit not small for eye-tracking studies specifically. Second, the questionnaire was quite lengthy. Third, better-educated people were overrepresented in the sample. These limitations were addressed in the second experiment.

In the second experiment, the findings on data quality were replicated with a larger, more balanced sample and a relatively short web survey. Again, even in the shorter survey, data quality was found to be lower for questions positioned later in a questionnaire. This is shown in less item differentiation, shorter response times, and differences in the response means.

The findings lead to some practical implications. Questionnaire designers and researchers should try to motivate their respondents to optimize when answering a questionnaire by keeping the questionnaires short and relevant for the respondents, bearing in mind that there might otherwise be a loss in data quality, especially towards the end of the survey. However, a certain amount of questions on a topic might be absolutely necessary. In these cases, one possibility in web surveys could be to rotate blocks of questions between

respondents, or to split the questionnaire into several short individual surveys (i.e., questionnaire modularization; Toepoel and Lugtig 2018). Being able to interrupt answering the questionnaire and resume it later is another strategy to allow respondents to be more attentive. Both studies presented here were self-administered online questionnaires. Worthwhile future research might investigate the effects of question positioning in postal surveys and in interviewer-administered surveys. By design, postal surveys allow the respondent to interrupt filling out the questionnaire and continue later. In interviewer-administered surveys, the interviewers play an important role in motivating respondents throughout questionnaire completion (Fowler 2013). Extending this research to other question types and formats would also be worthwhile.

It would also be interesting to know whether there is some kind of cut-off point at which motivation decreases and respondents stop answering conscientiously. However, the extent to which defining such a point in time would be universally applicable is unclear, as it probably depends on the individual interest in the survey topic as well as on the survey mode. Finally, it would be of interest to follow up on the finding made in the eye-tracking experiment that respondents who answered the grid question at a later position did not consider the items in the second half as thoughtfully as respondents who answered them earlier in the questionnaire while there was no statistically significant difference for the first half. As fixation time and fixation count are measures for attention, this might be an indication that attention not only decreases during the entire questionnaire but also within questions consisting of multiple items when placed later in the questionnaire. However, this interpretation should be treated with caution as the order of the items has not been varied in this study. Future studies manipulating not only the question position but also the order of the items are needed to further address this effect.

Author's Note

The SPSS data file containing the eye-tracking data is available by contacting the author. All data analyses were performed using SPSS Advances Statistics 24. For gaze analysis, the SMI BeGaze Version 3.6.57 was utilized.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

The supplemental material for this article is available online.

ORCID iD

Cornelia Eva Neuert  <https://orcid.org/0000-0001-9855-5618>

Note

1. Höhne and Schlosser (2018) examined the adequacy of response time outliers and concluded that the approach suggested by Hoaglin et al. (2000) is, in particular, distribution sensitive because the lower and upper quartile range is used as a measure of dispersion.

References

- Biemer, Paul P. and Lars Lyberg E. 2003. *Introduction to survey quality*. Hoboken, NJ: John Wiley.
- Burchell, Brendan and Catherine Marsh. 1992. "The Effect of Questionnaire Length on Survey Response." *Quality & Quantity* 26(3): 233-44.
- Callegaro, Mario and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72(5): 1008-32.
- Cannell, Charles F., Peter V. Miller, and Lois Oksenberg. 1981. "Research on Interviewing Techniques." *Sociological Methodology* 12(4): 389-437.
- Clancy, Kevin J. and Robert A. Wachsler. 1971. "Positional Effects in Shared-cost Surveys." *Public Opinion Quarterly* 35(2): 258-65.
- Craik, Fergus I. 1999. "Levels of Encoding and Retrieval." Pp. 97-104 in *Stratification in Cognition and Consciousness*, edited by Bradford H. Challis and Boris Velichkovsky. Amsterdam, the Netherlands: John Benjamins Publishing.
- Fowler, Floyd J. 2013. *Survey Research Methods* (5th ed.). Thousand Oaks, CA: Sage.
- Fowler, Floyd J. and Cannell Charles F. 1996. "Using Behavioral Coding to Identify Cognitive Problems with Survey Questions." Pp. 15-36 in *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by Norbert Schwarz and Sudman Schuman. San Francisco: Jossey-Bass.
- Galesic, Mirta. 2006. "Dropouts on the Web: Effects of Interest and Burden Experienced during an Online Survey." *Journal of Official Statistics* 22(2): 313-28.

- Galesic, Mirta and Michael Bosnjak. 2009. "Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey." *Public Opinion Quarterly* 73(2): 349-60.
- Groves, Robert M., Floyd Jr. Fowler, Mick P. Couper, James M. Lepkowski, Eleonor Singer, and Robert Tourangeau. 2009. *Survey Methodology* (2nd ed.). New York: Wiley.
- Groves, Robert, M., Eleonor Singer, Amy D. Corning, and Ashley Bowers. 1999. "A Laboratory Approach to Measuring the Effects on Survey Participation of Interview Length, Incentives, Differential Incentives, and Refusal Conversion." *Journal of Official Statistics* 15(2): 251-68.
- Galesic, Mirta, Roger Tourangeau, Mick P. Couper, and Fred G. Conrad. 2008. "Eye-tracking Data: New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding." *Public Opinion Quarterly* 72(5): 892-913.
- Galesic, Mirta and Ting Yan. 2011. "Use of Eye Tracking for Studying Survey Response Processes." Pp. 349-70 in *Social and Behavioral Research and the Internet: Advances in Applied methods and Research Strategies*, edited by Marcel Das, Peter Ester, and Lars Kaczmirek. New York: Routledge.
- Hansen, Kasper M. 2007. "The Effects of Incentives, Interview Length, and Interviewer Characteristics on Response Rates in a CATI-study." *International Journal of Public Opinion Research* 19(1): 112-21.
- Helgeson, James G. and Michael L. Ursic. 1994. "The Role of Affective and Cognitive Decision-making Processes during Questionnaire Completion." *Psychology & Marketing* 11(5): 493-510.
- Herzog, A. Regula and Jerald G. Bachmann. 1981. "Effects on Questionnaire Length on Response Quality." *Public Opinion Quarterly* 45(4): 549-59.
- Hoaglin, David C., Frederick Mosteller, and John Wilder Tukey. 2000. *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley.
- Höhne, Jan K. and Stephan Schlosser. 2018. "Investigating the Adequacy of Response Time Outlier Definitions in Computer-Based Web Surveys Using Paradata Survey Focus." *Social Science Computer Review* 36(3):369-78.
- Johnson, W. Russell, Nicholas A. Sieveking, and Earl S. Clanton. 1974. "Effects of Alternative Positioning of Open-ended Questions in Multiple-choice Questionnaires." *Journal of Applied Psychology* 59(6): 776-78.
- Joos, Markus, Matthias Rötting, and Boris M. Velichkovsky. 2003. "Die Bewegungen des menschlichen Auges: Fakten, Methoden, innovative Anwendungen." Pp. 142-68 in *Psycholinguistik/ Psycholinguistics. Ein internationales Handbuch/ An International Handbook*, edited by Gert Rickheit, Theo Herrmann, and Werner Deutsch. Berlin & NY: de Gruyter.

- Just, Marcel A. and Patricia A. Carpenter. 1980. "A Theory of Reading: From Eye Fixations to Comprehension." *Psychological Review* 87:329-54.
- Just, Marcel A. and Patricia A. Carpenter. 2018. "Using Eye Fixations to Study Reading Comprehension." Pp. 151-82 in *New Methods in Reading Comprehension Research*, edited by David E. Kieras and Marcel A. Just. Abingdon: Routledge.
- Kaczmirek, Lars. 2005. A Framework for the Collection of Universal Client Side Paradata (UCSP). Webpage: <http://kaczmirek.de/ucsp/ucsp.html> (last update November 19, 2019).
- Kovaleva, A., C. Beierlein, C. J. Kemper, and B. Rammstedt. . 2014. "Die Skala Impulsives-Verhalten-8 (I-8)." *Zusammenstellung sozialwissenschaftlicher Items und Skalen*.
- Krosnick, Jon, A., Allyson L. Holbrook, Matthew K. Berent, Richard T. Carson, W. Michael Hanemann, Raymond J. Kopp, Robert Cameron Mitchell, Stanley Presser, Paul A. Ruud, V. Kerry Smith, Wendy R. Moody, Melanie C. Green, and Michael Conaway. 2002. "The Impact of "No Opinion" Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66(3): 371-403.
- Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5(3): 213-236.
- Krosnick, Jon A. and Duane Alwin. 1987. "An Evaluation of a Cognitive Theory of Response-order Effects in Survey Measurement." *Public Opinion Quarterly* 51(2):201-19.
- Krosnick, Jon A. and Duane F. Alwin 1988. A Test of the Form-Resistant Correlation Hypothesis. *Public Opinion Quarterly* 52:526-538.
- Peytchev, Andy and Emilia Peytcheva. 2017. "Reduction of Measurement Error Due to Survey Length: Evaluation of the Split Questionnaire Design Approach." *Survey Research Methods* 11(4): 361-68.
- Revilla, Melanie and Carlos Ochoa. 2017. "Ideal and Maximum Length for a Web Survey." *International Journal of Market Research* 59(5): 557-65.
- Romano Bergstrom, Jennifer R. and Andrew Schall. 2014. "Introduction to Eye Tracking." Pp. 3-46 in *Eye Tracking in user Experience Design*, edited by J. R. Bergstrom and Andrew Schall. San Francisco, CA: Morgan Kaufmann.
- Schober, Michael F. and Frederick G. Conrad. 1997. "Does Conversational Interviewing Reduce Survey Measurement Error?" *Public Opinion Quarterly* 61(4): 576-602.
- Schwarz, Norbert and Hippler Hans-J. 1991. Response alternatives: The impact of their choice and presentation order. In Paul P. Biemer, Robert M. Groves, Lars E.

- Lyberg, Nancy A. Mathiowetz, and Seymour Sudman (Eds.), *Measurement errors in surveys* (pp. 41-56). Hoboken, NJ, USA: John Wiley & Sons.
- Shields, Jennifer and Nhien To. 2005. "Learning to say No: Conditioned Underreporting in an Expenditure Survey." Pp. 3963-68 in *American Association for Public Opinion Research-American Statistical Association, Proceedings of the Section on Survey Research Methods*.
- Singer, Eleanor, John Van Hoewyk, Nancy Gebler, Trivellore Raghunathan, and Katherine McGonagle. 1999. "The Effect of Incentives on Response Rates in Interviewer-mediated Surveys." *Journal of Official Statistics* 15(2): 217-30.
- Sudman, Seymour and Norman M. Bradburn. 1974. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Toepoel, Vera and Peter Lugtig. 2018. "Modularization in an Era of Mobile Web: Investigating the Effects of Cutting a Survey into Smaller Pieces on Data Quality." *Social Science Computer Review*. <https://doi.org/10.1177/0894439318784882>
- Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Whiteside, Stephen P. and Donald R. Lynam. 2001. "The Five Factor Model and Impulsivity: Using a Structural Model of Personality to Understand Impulsivity." *Personality and Individual Differences* 30(4): 669-89.
- Zhang, Chan and Frederick Conrad. 2014. "Speeding in Web Surveys: The Tendency to Answer Very Fast and its Association with Straightlining." *Survey Research Methods* 8(2): 127-35.

Author Biography

Cornelia Eva Neuert is a postdoctoral researcher and team leader at GESIS-Leibniz Institute for the Social Sciences. Her research interests include questionnaire design, questionnaire evaluation, and data quality. Her research has been published in *Field Methods*, *Journal of Survey Statistics and Methodology*, *Social Science Computer Review*, and *International Journal of Social Research Methodology*.