

# Examining final comment questions with requests for written and oral answers

International Journal of  
Market Research  
2024, Vol. 66(5) 550–558  
© The Author(s) 2024



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/14707853241229329  
[journals.sagepub.com/home/mre](https://journals.sagepub.com/home/mre)



**Jan Karem Höhne** 

German Centre for Higher Education Research and Science Studies; Leibniz University Hannover, Germany

**Joshua Claassen** 

German Centre for Higher Education Research and Science Studies; Leibniz University Hannover, Germany

## Abstract

Self-administered web surveys provide respondents only limited opportunities for feedback. Therefore, many web surveys include so-called “final comment questions” (FCQs) that allow respondents to elaborate on the survey in general and the survey questions in particular. Typing in open answers in text boxes is burdensome – especially via smartphones with virtual on-screen keypads – so that respondents frequently provide short or no answers at all. In this study, we make use of new technological advancements in web survey methodology and investigate FCQs with requests for written and oral answers. For this purpose, we conducted an experiment in a smartphone survey ( $N = 1,001$ ) in a German online panel. The results reveal that FCQs with a request for written and oral answers do not differ with respect to item-nonresponse. However, oral answers are substantially longer than written answers pointing to more in-depth information. The oral answer condition also results in more positive comments than the written answer condition. This study is a methodological showcase for innovative web survey design that contributes to the improvement of data quality.

## Keywords

answer behavior, built-in microphone, final comment question, open-ended answers, smartphone survey, survey evaluation

---

## Corresponding author:

Joshua Claassen, Research Infrastructure and Methods Division, German Centre for Higher Education Research and Science Studies, Lange Laube 12, Hannover 30159, Germany.

Email: [claassen@dzhw.eu](mailto:claassen@dzhw.eu)

---

**Correction (February 2024):** This article has been updated with the corresponding author’s affiliation and formatting correction in the Tables.

Data Availability Statement included at the end of the article.

## Introduction and background

Self-administered web surveys are a prevailing data collection method in behavioral and social science research. Typically, such web surveys include questions with closed answer formats that restrict respondents' answers to a pre-defined list of options (e.g., running from "agree strongly" to "disagree strongly"). Despite the prevalence of closed answer formats researchers increasingly employ questions with open answer formats (e.g., text fields for typing in answers) because they facilitate the collection of in-depth and unfiltered information (Revilla & Ochoa, 2016; Smyth et al., 2009). However, respondents may find it difficult to articulate themselves in writing and to type in answers; especially through virtual on-screen keypads of smartphones that shrink the viewing space (Höhne et al., 2020). Thus, open questions frequently result in very short or no answers at all indicating that respondents may not reveal all their thoughts because of the answer provision process (Revilla & Ochoa, 2016).

Another methodological shortcoming of self-administered web surveys is the absence of interviewers limiting respondents' chance to provide feedback on the web survey in general and its questions in particular. Thus, researchers frequently implement so-called "final comment questions" (FCQs), as a special type of open questions, at the end of the web survey. One example of a FCQ is the following question: "Do you have any further comments on the survey or its questions?". FCQs usually include an open answer format in the form of one or multiple text boxes requesting respondents to type in answers in their own words (Decorte et al., 2019; McLauchlan & Schonlau, 2016; O'Cathain & Thomas, 2004; Schonlau, 2015). They can shed light on general reactions (e.g., respondents' feelings and thoughts), topic coverage (e.g., identifying missing aspects that were not included in the survey), potential critique (e.g., irritating or overly sensitive questions), methodological problems (e.g., missing answer options or question redundancy), and technological issues (e.g., dysfunctional survey operation or navigation issues). Thus, they contribute to the improvement of web survey design and help to evaluate the quality of respondents' answers.

New technological developments, coupled with a continuous increase in web surveys completed via smartphones (Gummer et al., 2023), extend the methodological toolkit for collecting respondents' answers to open questions in general and FCQs in particular. Specifically, the built-in microphones of smartphones facilitate the administration of open questions with requests for oral instead of written answers (Gavras et al., 2022; Gavras & Höhne, 2022; Revilla & Couper, 2021; Revilla et al., 2020). The use of oral input, compared to written input, has the great potential to ease the answer provision process, as in the former case, respondents only need to press a recording button and record their answers (Gavras & Höhne, 2022). Gavras et al. (2022) argue that requests for written answers are more likely to trigger a memory-based processing (Tourangeau et al., 2000; Zaller & Feldman, 1992) and may thus be more intentional and conscious, whereas oral answers are more likely to trigger an on-line processing (Lodge et al., 1989; McGraw et al., 2003) and may thus be more intuitive and spontaneous. The presumably less burdensome answer delivery associated with oral answers, coupled with an on-line processing, potentially results in longer answers. In line with this reasoning, research has shown that oral answers are longer and consist of more topics than their written counterparts (Gavras et al., 2022). Topics also vary between both answer conditions. Oral answers also perform better than written ones when it comes to validity (Gavras & Höhne, 2022). However, item-nonresponse appears to be higher in oral than written answers (Revilla & Couper, 2021; Revilla et al., 2020). This indicates that a substantial minority of respondents is not willing and/or able to provide oral answers.

To our best of knowledge, no previous studies have compared respondents' answers to FCQs using requests for written and oral answers. In order to fill this research gap, we conducted an

experiment in a smartphone survey and randomly assigned respondents to different answer request conditions. Some respondents were asked to provide written answers and some others were asked to provide oral answers. In line with the existing research on written and oral answers in smartphone surveys, we address the following three research questions (RQs):

Do FCQs with requests for written and oral answers differ in terms of

... *item-nonresponse?* (RQ1)

... *answer length?* (RQ2)

... *content of comments?* (RQ3)

## Method

### Data collection

Data were collected in the Forsa Omninet Panel ([omninet.forsa.de](https://omninet.forsa.de)) in Germany in November 2021. Forsa drew a cross-quota sample from their online panel based on age (young, middle, and old) and gender (female and male). In addition, they drew quotas on education (low, medium, and high). The quotas were calculated based on the German Microcensus, which served as a population benchmark. Data and analysis code for replication purposes are available through Harvard Dataverse (see <https://doi.org/10.7910/DVN/J8K7VW>).

The email invitation included information on the device to be used for survey participation (smartphone) and a link that re-directed respondents to the web survey. The first web survey page introduced the topic and outlined the overall procedure. In addition, it included a statement of confidentiality assuring that the study adheres to existing data protection laws and regulations. Prior informed consent for data collection was obtained by Forsa. Respondents also received financial compensation for their participation from Forsa.

Respondents were randomly assigned to one out of two experimental groups. The first experimental group received one FCQ with a request for a written answer (written condition). The second experimental group received one identical FCQ but with a request for an oral answer (oral condition).

In order to collect oral answers, we used the open-source “SurveyVoice” (SVoice) tool programmed by Hühne et al. (2021). SVoice resembles the voice recording function of popular instant messengers, such as WhatsApp and WeChat, and works for both Android and iOS devices. It is browser-based and does not require respondents to download any software or plugins.

### Sample

Forsa invited 6,745 respondents to take part in the web survey. No respondents were screened out because of full quotas or because they tried to access the web survey with another device than a smartphone. A total of 1,681 respondents started the web survey, but 680 of them broke-off before they were asked any study-relevant questions. In the written condition 159 (about 24%) respondents broke-off, whereas in the oral condition 521 (about 51%) respondents broke-off. In this study, we focus on respondents that finished the entire smartphone survey (full completes).

Of the 1,001 respondents, 500 took part in the written condition and 501 took part in the oral condition. Participation rate was about 15% among all invitees. Table 1 shows the sample characteristics of the written and oral conditions.

**Table 1.** Sample characteristics of the written and oral conditions.

Respondent characteristics	Written condition	Oral condition
Age	48	49
Gender: Female	50	48
Education: Medium	41	43
Education: High	29	26
Smartphone skills	5.6	5.6
Internet usage	6.1	6.0

Note. We report means for age (in years), smartphone skills (1 “Very bad” to 7 “Very good”), and Internet usage (1 “Not at all” to 7 “Very often”). For gender (0 “Male” and 1 “Female”) and education (1 “Low,” 2 “Medium,” and 3 “High”), we report percentages instead.

In order to evaluate the effectiveness of random assignment, we compared the sample composition between the written and oral conditions. We found no statistically significant differences with respect to age, gender, education, smartphone skills, and Internet usage.

### *Final comment question*

We asked respondents the following FCQ at the end of the web survey:

Finally, we would like to give you the opportunity to say something about our survey. Do you have any comments or suggestions on the survey in general or individual questions in particular?

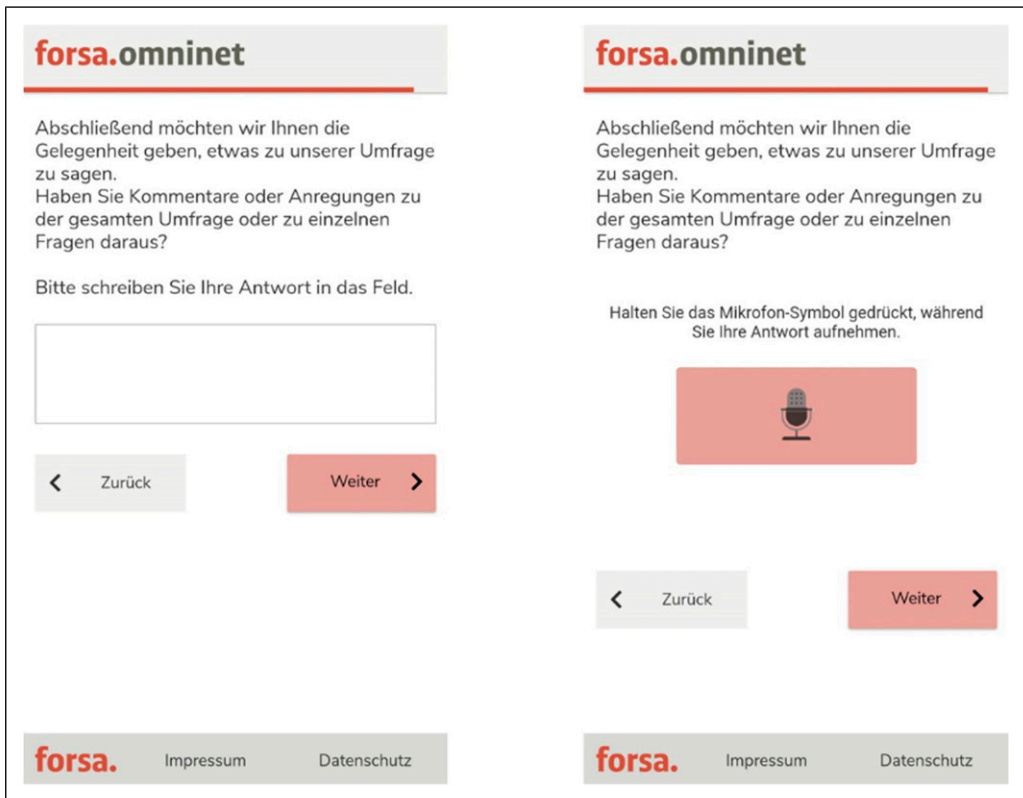
Depending on the experimental condition, respondents were either asked to provide a written answer or to provide an oral answer. [Figure 1](#) provides exemplary screenshots of the FCQs.

At the beginning of the web survey, respondents received a description on how to provide text or voice answers (the online [Appendix](#) provides English translations of these descriptions). In total, the survey consisted of eight open questions with requests for written or oral answers that dealt with a variety of topics, such as the German vaccination campaign and working women. These questions preceded the FCQ.

## **Results**

In order to investigate our first research question, we compared item-nonresponse (i.e., the percentage of respondents that provided no answer to the FCQ) between the written and oral request conditions. With less than 50% of respondents providing an answer in both conditions item-nonresponse is relatively high. In the written condition the item-nonresponse rate is 55.8%, whereas in the oral condition the item-nonresponse rate is 52.3%. We conducted a Z-test for determining statistical significance using a significance level of 0.05 as threshold. The results show no statistically significant difference between the conditions with requests for written and oral answers ( $Z = 1.11, p = .27$ ).

With respect to our second research question, we compared respondents’ answer length (i.e., the number of words that respondents provided).<sup>1</sup> For this purpose, we excluded all respondents that engaged in item-nonresponse from statistical analyses. The median answer length is 8 words in the written condition and 27 words in the oral condition. Thus, oral answers are more than three times longer than their written counterparts. The result of a non-parametric median test shows a statistically significant difference between both request conditions ( $\chi^2(1) = 62.22, p = .00$ ).



**Figure 1.** Screenshots of the FCQs with requests for written and oral answers.

Note. FCQ with a request for a written answer on the left and FCQ with a request for an oral answer on the right.

Finally, in order to address our third research question, we investigate respondents' answers to the FCQ with written and oral answer requests. For this purpose, we manually coded respondents' answers with the help of a student assistant building on the coding scheme by [Schonlau \(2015\)](#) with 9 categories. During the coding, we partially revised and expanded the coding scheme so that it finally consisted of 13 categories (i.e., 4 positive, 4 neutral, 4 negative, and 1 nonsense). Using this final coding scheme, the second author independently coded a randomly selected subset of about 30% ( $n = 138$ ) of the answers (excluding answers that were coded as nonsense). A Cohen's kappa value of 0.84 shows a high agreement between the two coders ([Landis & Koch, 1977](#)). Importantly, comments were coded into one single category. If two categories applied to one answer, the category corresponding to the first mentioned content in the answer was selected. Answer analysis remains on a descriptive level by reporting percentages.

Starting with positive comments [Table 2](#) shows that in both conditions most respondents commented on open questions and the answer delivery process (e.g., "I thought it was great that you could comment on the questions in such detail"). This was twice as high in the oral condition than in the written condition. In the oral condition, about 11% stated that the survey was fun, interesting, and good (only about 5% did so in the written condition). Some few respondents (less than 3%)

**Table 2.** Categorization of respondents' answers to the FCQ.

Content of comments	Written condition (%)	Oral condition (%)
<i>Positive</i>		
Open questions and answer provision process	10.4	20.9
Fun, interesting, and good	5.4	11.3
Results release	2.3	0.4
Other	6.8	3.4
<i>Neutral</i>		
Open questions and answer provision process	3.6	13.0
Personal views and answer elaboration	6.3	9.2
General or non-informative comment	33.0	20.5
Other	5.9	2.9
<i>Negative</i>		
Open questions and answer provision process	10.9	11.7
Smartphone use for survey completion	5.4	0.4
Unclear questions	4.1	2.9
Other	5.0	1.7
<i>Nonsense</i>	0.9	1.7

asked for results release and about 7% (written condition) and 3% (oral condition) provided other comments (e.g., "Better than the previous survey").

With respect to neutral comments, we found that most respondents provided a general (e.g., "Everything was all right") or non-informative comment (e.g., "I have no comments to make"). In the written condition this is higher (33%) than in the oral condition (21%). About 4% (written condition) and 13% (oral condition) commented on open questions and the answer delivery process (e.g., "I wonder how you analyze that" and "This was a somewhat unusual survey for me with this sort of format"). Between 9% and 3% of respondents provided personal views and answer elaborations (e.g., "I think it would be very important if politicians and citizens would govern together again") or other comments (e.g., "What was the purpose of this survey?").

When it comes to negative comments about 11% (written condition) and 12% (oral condition) complained about open questions and the answer provision process (e.g., "It is more difficult to put something into words than to choose from given answers"). Another 5% (written condition) complained about smartphone use for survey completion (e.g., "Next time, please also include the option of using the PC or tablet"). In the oral condition, this was done by less than 1% of the respondents. Less than 5% of respondents (both written and oral conditions) complained about unclear questions (e.g., "Would have been nice if the questions were a bit more specific"). Finally, about 5% (written condition) and 2% (oral condition) of respondents provided other comments (e.g., "The topics were of little interest to me").

Some of the respondents also gave nonsense answers (e.g., "War"). However, these answers account for less than 2% of the answers in both conditions.

## Discussion and conclusion

The goal of this study was to investigate "final comment questions" (FCQs) with requests for written and oral answers. For this purpose, we conducted an experiment in a smartphone survey and randomly assigned respondents to one out of two request conditions. The overall results reveal no

differences with respect to item-nonresponse, but substantial differences with respect to answer length. The content of respondents' answers to the FCQ also differs.

Previous studies on open questions with requests for written and oral answers consistently found higher item-nonresponse for oral answers (Gavras et al., 2022; Gavras & Höhne, 2022; Revilla & Couper, 2021; Revilla et al., 2020). In this study, item-nonresponse does not differ significantly between the two request conditions. Item-nonresponse is even slightly higher in the written than in the oral condition. This finding indicates that requests for oral answers may encourage respondents to provide answers to FCQs. However, this conjecture lacks empirical evidence and thus it needs further, more refined investigation.

The length differences between written and oral answers indicate differences in response burden. Written answers appear to be more burdensome than oral answers. Most likely, the answer delivery process is responsible for the length differences. Entering written answers via a virtual on-screen keypad of smartphones is more burdensome than recording an oral answer by simply pressing a recording button (Gavras et al., 2022). In addition, this finding indicates that oral answers, compared to their written counterparts, trigger open narrations that potentially result in more nuanced and in-depth information. We see the underlying answer processes responsible for this finding. While written answers seem to initiate an intentional and conscious memory-based processing, oral answers seem to initiate an intuitive and spontaneous on-line processing. In order to go beyond answer length, it would be worthwhile to compare the number of topics between both request conditions using Structural Topic Models (STM; Roberts et al., 2014). This potentially helps to draw more refined conclusions about the informative content of written and oral answers.

The coding of respondents' answers to the FCQs resulted in four overall categories: positive, neutral, negative, and nonsense. The oral request condition produced more positive comments than its written counterpart. Providing oral answers in a smartphone survey seems to appeal to a substantial minority of respondents.

This study has some limitations that provide perspectives for future research. First, we drew a quota sample from a nonprobability online panel. In order to draw more robust conclusions, it would be worthwhile to rerun a similar study in a probability-based panel. Relatedly, it would be worthwhile to conduct subgroup analyses that, for example, take demographic characteristics, such as age and education, into consideration. This would help to draw a full picture of the usefulness and usability of oral answers when it comes to FCQs. Second, an item-nonresponse rate of 50% is high and thus it would be worthwhile to conduct research on how to decrease item-nonresponse associated with FCQs in general. For example, future studies may investigate the impact of motivational prompts on item-nonresponse and/or more refined incentive strategies. Specifically, it might be worthwhile to incentivize respondents per (oral or written) answer that they provide instead of providing them with a fixed, overall incentive.

FCQs with requests for oral answers provide some methodological advantages in smartphone surveys. They result in a comparable item-nonresponse rate as their written counterparts. This was not observed in previous studies on general open questions. In addition, they provide longer answers pointing to open narrations that potentially lead to more information from respondents. Respondents' comments are also more positive, which may enhance survey satisfaction. Thus, FCQs with requests for oral answers add to the methodological toolkit and warrant further research.

## **Acknowledgements**

The authors are grateful to Jessica Kuhlmann (Ruhr University Bochum) for her great support during data coding.

## 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.

## Ethical statement

### *Ethical approval*

The study presented in the manuscript was conducted in accordance with established ethical standards. Furthermore, this manuscript has not been published or accepted for publication elsewhere nor is it under consideration at another journal.

## ORCID iDs

Jan Karem Höhne  <https://orcid.org/0000-0003-1467-1975>

Joshua Claassen  <https://orcid.org/0009-0002-5492-4439>

## Data Availability Statement

Data and analysis code for replication purposes are available through Harvard Dataverse (see <https://doi.org/10.7910/DVN/J8K7VW>).

## Supplemental Material

Supplemental material for this article is available online.

## Note

1. Before data analysis, the recordings of respondents' oral answers were automatically transcribed by OpenAI's automatic speech recognition system Whisper (Radford et al., 2023). As a quality assurance measure, a student assistant listened to 20% of the recordings ( $n = 48$ ) and systematically notated any differences between the recordings and the transcripts. The differences were assessed by the second author, revealing only minor discrepancies and an overall high transcription quality. In addition, the second author manually examined and, if necessary, corrected the transcriptions of audio files with a duration of less than 2 seconds as Whisper sometimes encountered problems with extremely short audio files.

## References

- Decorte, T., Malm, A., Sznitman, S. R., Hakkarainen, P., Barratt, M. J., Potter, G. R., Werse, B., Kamphausen, G., Lenton, S., & Asmussen Frank, V. (2019). The challenges and benefits of analyzing feedback comments in surveys: Lessons from a cross-national online survey of small-scale cannabis growers. *Methodological Innovations*, 12(1), 205979911982560. <https://doi.org/10.1177/2059799119825606>
- Gavras, K., & Höhne, J. K. (2022). Evaluating political parties: Criterion validity of open questions with requests for text and voice answers. *International Journal of Social Research Methodology*, 25(1), 135–141. <https://doi.org/10.1080/13645579.2020.1860279>
- Gavras, K., Höhne, J. K., Blom, A. G., & Schoen, H. (2022). Innovating the collection of open-ended answers: The linguistic and content characteristics of written and oral answers to political attitude questions. *Journal of the Royal Statistical Society*, 185(3), 872–890. <https://doi.org/10.1111/rssa.12807>



- Gummer, T., Höhne, J. K., Rettig, T., Roßmann, J., & Kummerow, M. (2023). Is there a growing use of mobile devices in web surveys? Evidence from 128 web surveys in Germany. *Quality and Quantity*, 57(6), 5333–5353. <https://doi.org/10.1007/s11135-022-01601-8>
- Höhne, J. K., Gavras, K., & Qureshi, D. D. (2021). *SurveyVoice (SVoice): A comprehensive guide for collecting voice answers in surveys*. Zenodo. <https://doi.org/10.5281/zenodo.4644590>
- Höhne, J. K., Schlosser, S., Couper, M. P., & Blom, A. (2020). Switching away: Exploring on-device media multitasking in web surveys. *Computers in Human Behavior*, 111, Article 106417. <https://doi.org/10.1016/j.chb.2020.106417>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lodge, M., McGraw, K. M., & Stroh, P. (1989). An impression-driven model of candidate evaluation. *American Political Science Review*, 83(2), 399–419. <https://doi.org/10.2307/1962397>
- McGraw, K. M., Hasecke, E., & Conger, K. (2003). Ambivalence, uncertainty, and processes of candidate evaluation. *Political Psychology*, 24(3), 421–448. <https://doi.org/10.1111/0162-895X.00335>
- McLauchlan, C., & Schonlau, M. (2016). Are final comments in web survey panels associated with next-wave attrition? *Survey Research Methods*, 10(3), 211–224. <https://doi.org/10.18148/srm/2016.v10i3.6217>
- O’Cathain, A., & Thomas, K. J. (2004). Any other comments?’. Open questions on questionnaires – a bane or bonus to research?’. *BMC Medical Research Methodology*, 4, 25. <https://doi.org/10.1186/1471-2288-4-25>
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). Robust speech recognition via large-scale weak supervision. Proceedings of the 40th International Conference on Machine Learning, Honolulu Hawaii, USA, 23–29 July 2023, (pp. 28492–28518). <https://dl.acm.org/doi/10.5555/3618408.3619590>
- Revilla, M., & Couper, M. P. (2021). Improving the use of voice recording in a smartphone survey. *Social Science Computer Review*, 39(6), 1159–1178. <https://doi.org/10.1177/0894439319888708>
- Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). Testing the use of voice input in a smartphone web survey. *Social Science Computer Review*, 38(2), 207–224. <https://doi.org/10.1177/0894439318810715>
- Revilla, M., & Ochoa, C. (2016). Open narrative questions in PC and smartphones: Is the device playing a role? *Quality and Quantity*, 50(6), 2495–2513. <https://doi.org/10.1007/s11135-015-0273-2>
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082. <https://doi.org/10.1111/ajps.12103>
- Schonlau, M. (2015). What do web survey panel respondents answer when asked ‘Do you have any other comment?’. In *Survey methods: Insights from the field*. FORS - Swiss Foundation for Research in Social Sciences, GESIS Leibniz-Institute for the Social Sciences, and PUMA - Platform for Surveys, Methods and Empirical Analyses. <https://doi.org/10.13094/SMIF-2015-00013>
- Smyth, J. D., Dillman, D. A., Christian, L. M., & McBride, M. (2009). Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly*, 73(2), 325–337. <https://doi.org/10.1093/poq/nfp029>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American Journal of Political Science*, 36(3), 579–616. <https://doi.org/10.2307/2111583>