Real-time prompts: Investigating the efficacy of speeding feedback in a web survey

Jan Karem Höhne
DZHW
Leibniz University Hannover

Achim Goerres
University of Duisburg-Essen

Joshua Claassen
DZHW
Leibniz University Hannover

Abstract

Self-administered web surveys are omnipresent and can be speculated to be by far the dominant survey mode. However, respondents often answer such surveys extremely fast (called "speeding") and do not carefully read, process, and answer questions. To reduce respondent speeding, researchers can incorporate real-time feedback that urges respondents to slow down and provide careful answers. However, research on speeding feedback and its optimal visual design is scarce. To address this gap in the literature, we conducted a web survey (n = 2,006) in the Forsa Omninet Panel in Germany and randomly assigned respondents to speeding feedback with no visual cue, a neutral cue (image of a paperclip), or a humanized cue (photo of the principal investigator). By estimating multilevel regressions, we investigate how speeding feedback with different visual cues affects speeding and data quality in later questions. The results reveal that speeding feedback successfully reduces speeding in subsequent questions, irrespective of the visual cue. However, the results on the relation between speeding feedback and data quality are somewhat mixed and partially depend on survey question characteristics.

Keywords: Answer behavior, Data quality, Paradata, Real-time feedback, Satisficing, Survey prompts

Introduction and research questions

During the last decade, survey data collection has fundamentally changed. Cost-efficient and streamlined web surveys successively replace other survey modes, especially in-person interviews (Schober, 2018). Even many established, large-scale surveys using in-person interviews, such as the American National Election Study (ANES), the European Social Survey (ESS), and the Health and Retirement Study (HRS), have started to explore data collection methods that involve web surveys. This transition towards web surveys was further accelerated by the global Covid-19 pandemic from 2020 to 2022 in which public regulations, coupled with ethical concerns, made in-person interviews impossible (Saarijärvi & Bratt, 2021; Self, 2021).

The key problem associated with this transition is that web surveys may not be entirely ready for taking over as the dominant survey mode. The literature includes numerous examples

---

This document is a preprint and thus it may differ from the final version.

of superficial answer behavior observed in web surveys compromising data quality. For example, Zhang and Conrad (2014) show that extremely fast answering without the possibility to carefully read and process the survey question under consideration (called "speeding") is associated with superficial answers in the form of non-differentiation (i.e., selecting the same answer option across questions; Krosnick, 1991). Similarly, Malhotra (2009) provides evidence that speeding respondents are more prone to primacy effects (i.e., selecting the first answer option of the scale; Sudman et al., 1996). Speeders may be driven by finishing web surveys quickly, possibly enticed by the incentive that is often paid for web survey completion, and not by answering thoughtfully.

In contrast to in-person interviews, self-administered web surveys do not have interviewers to motivate respondents and to monitor their answer behavior. It is up to respondents to remain conscientious during web survey participation. One way to compensate for the absence of interviewers and to promote high-quality answers is to incorporate interactive tools in web surveys. These tools automatically and unobtrusively monitor respondents' answer behavior throughout the web survey without any human supervision. If respondents exhibit superficial answer behavior, such as speeding, these tools flag the particular behavior.

Recently, researchers have started to incorporate real-time feedback in web surveys to improve answer behavior and data quality. Such feedback is based on respondents' paradata (i.e., automated data providing information about the answering process that can be used to describe and evaluate this process; Couper, 2000). Certain types of paradata indicate that respondents do not pay close attention to the web survey (Conrad, Tourangeau, Couper, & Zhang, 2017; Höhne, Schlosser, Couper, & Blom, 2020; Sendelbah, Vehovar, Slavec, & Petrovčič, 2016; Zhang & Conrad, 2018). This includes response times that, for example, inform about speeding and therefore suggest insufficient thought before answering (Conrad et al., 2017; Zhang & Conrad, 2018). If respondents engage in speeding, they are notified that their care and sincerity while answering matter. Potentially, this ensures that respondents stay motivated and conscientious, thereby stimulating accurate and informative answers.

Feedback on respondents' answer behavior is not widely used in web surveys yet. The study by Conrad et al. (2017) stands out of the existing literature. The authors provided respondents feedback when they answered faster than a minimal response time threshold based on typical reading speeds. In six web survey experiments, such feedback has proven its worth by slowing respondents down and enhancing the quality of their answers. This was found regardless of whether the speeding feedback occurred early or late in the web survey, in the first or later waves of a longitudinal web survey, among participants from non-probability or probability-based online panels, and whether the feedback was provided only on the first or all speeding occasions. Feedback did not increase missing data in terms of dropouts, indicating that respondents were not annoyed to such an extent as to drop out. However, feedback on speeding decreased non-differentiated answers on later questions, suggesting a lasting data quality improvement.

Interestingly, the type of visual cues implemented in feedback may also matter. Using data from a non-probability online panel, Zhang and Conrad (2018) show that feedback with a humanized cue (photo of a call center operator), compared to feedback with a computer-like cue (yellow triangle error sign), resulted in somewhat less speeding. Speeding was highest in the control group without any visual cues. The disclosure of sensitive information was higher when speeding feedback was accompanied by no visual cue or a humanized cue. These findings suggest that humanized cues in web surveys may positively affect respondents' answer behavior, leading to less speeding, and higher data quality.

So far, few studies have been carried out on speeding feedback in web surveys (in contrast to research on attention checks; see Shamon & Berning, 2020). This especially applies to studies that investigate the efficacy of speeding feedback in relation to humanized cues, such as a photo of the principal investigator (see, for example, Tourangeau et al., Couper, & Steiger, 2003). Speeding feedback accompanied by humanized cues may increase social presence by another person, potentially leading respondents to feel more obliged to follow instructions (Zhang & Conrad, 2018). To put it differently, the creation of social presence through the employment of photos of humans may increase the efficacy of speeding feedback. This line of research has its origin in human-computer interaction research (Nass, Moon, & Carney, 1999; Nass, Moon, & Green, 1997; Reeves & Nass, 1997). Nass and colleagues, for example, argue that people treat computerized systems – including web surveys – as social actors rather than lifeless instruments (see also Tourangeau et al., 2003, p. 2). This especially applies when computerized systems consist of humanized cues, such as human voice and embodied human attributes.

Since these previous studies, little has been done on speeding feedback in web surveys. Particularly, the investigation of visual cues embedded in speeding feedback is scarce so that the impact on speeding and data quality remains wide open. Intriguing examples are the studies by Conrad et al. (2017) and Zhang and Conrad (2018). To the best of our knowledge, there is no research on respondents' disposition towards or perception of speeding feedback. In this study, we therefore attempt to close this research gap and to shed light on the efficacy of speeding feedback, coupled with visual cues. To this end, we address the following three research questions (RQs):

RQ1: How does speeding feedback with different visual cues affect speeding?
RQ2: How does speeding feedback with different visual cues affect data quality?
RQ3: How is speeding feedback with different visual cues evaluated by respondents?

We conducted a web survey in the Forsa Omninet Panel in Germany providing speeding feedback to respondents while answering 15 closed questions on political solidarities and associated concepts. Respondents were randomized to different feedback groups varying visual cues. We also asked speeding respondents to answer several evaluative closed questions on how they perceived the feedback. In doing so, our study stands out from previous research advancing the current state of research.

In what follows, we describe the data collection, sample characteristics, and experimental design, as well as the questions and speeding feedback employed in this study. We then outline our analytical strategy and report our statistical results. Finally, we discuss our empirical findings and provide avenues for future research.

Method

Data collection

Data were collected in the Forsa Omninet Panel (omninet.forsa.de) in Germany in July 2023. While the survey mode of the Omninet Panel is online, respondents are recruited offline through a probability-based telephone sample. Respondents could not sign up themselves preventing mock accounts and duplicates. Forsa drew a cross-quota sample from their panel based on age (young, middle, and old) and gender (female and male). We also included quotas on education (low, middle, and high) and region (eastern and western Germany). The quotas were calculated based on the German Microcensus, which served as a population benchmark.

Forsa invited respondents via email (including two rounds of reminders). The email informed respondents that they would participate in a web survey conducted by the University of Duisburg-Essen. In addition, it included a link directing respondents to the web survey. On the first page of the web survey, respondents were introduced to the topic (i.e. social and political attitudes) and the procedure of the web survey. Respondents also received a statement of confidentiality assuring them that the study adheres to existing data protection laws and regulations.

As to research ethics, the study was pre-approved by the ethics committee of the department of Computer Science and Applied Cognitive Science of the University of Duisburg-Essen. Respondents received modest financial compensation for their participation from Forsa. Similar to previous studies (Conrad et al., 2017; Zhang & Conrad, 2018) we set a minimal response time threshold to determine speeding on a question level: 300 milliseconds (msec) per word (Zhang & Conrad, 2018). For example, the speeding threshold for a ten-word question would have been 3,000 msec. In this study, the speeding thresholds varied between 3,000 msec and 22,800 msec. Importantly, we have divided the 15 closed questions for which we provided speeding feedback into five blocks (three closed questions per block). If respondents answered faster than the previously set threshold for at least one question out of the three questions in the block, they received speeding feedback at the end of the respective block. Thus, respondents could receive a maximum of five speeding prompts. Figure 1 illustrates the scheme for providing speeding feedback.

We collected response times using the open-source tool "Embedded Client Side Paradata (ECSP)" programmed by Schlosser and Höhne (2018, 2020). Prior informed consent for the collection of paradata was obtained by Forsa as part of the respondents' registration process.

Sample characteristics

Forsa invited a total of 5,200 respondents to participate in the web survey, of which 2,501 (48%) did not react to the survey invitation, 204 (4%) were screened out because quotas were already achieved, and 489 (9%) did not finish the web survey. This leaves us with 2,006 respondents available for statistical analyses (participation rate of about 39%).

These respondents were aged between 18 and 89 years, with a mean age of 53 years (standard deviation of 16 years), and 50% of them were female. In terms of education, 43% completed lower secondary school or less (low education level), 24% intermediate secondary school (medium education level), and 33% college preparatory secondary school or university (high education level). Overall, 44% of the respondents participated with a computer, 3% with a tablet, and 52% with a smartphone.

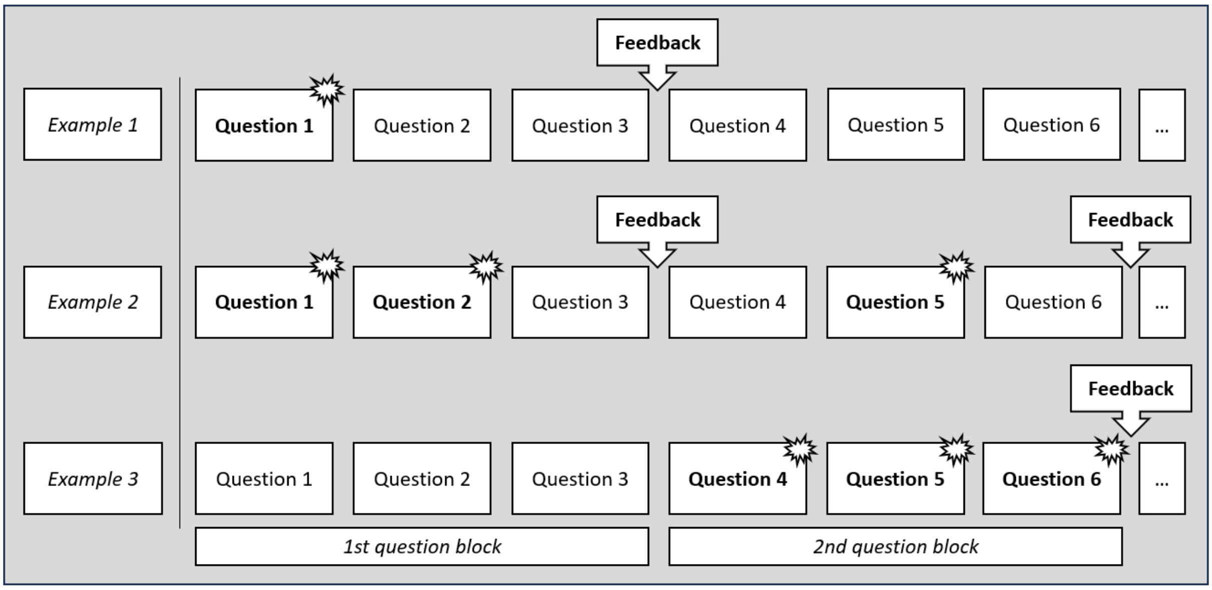

Figure 1. Three examples of the scheme for providing speeding feedback

Note. The burst icons in the right upper corner of the rectangles indicate that the question was answered faster than the set speeding threshold and that speeding feedback was provided at the end of the respective question block.

Experimental design

In this study, we used a between-subject design. Respondents were randomly assigned to one of three experimental groups. Table 1 describes these groups.

Table 1. Description of the experimental groups

Experimental group	Speeding feedback	Group size
1	No visual cue (control)	663
2	Neutral cue (paper clip)	689
3	Humanized cue (photo of PI)	654

Note. PI stands for principal investigator. We used a photo of the second author.

To evaluate the effectiveness of random assignment and the sample composition between the three experimental groups, we conducted several statistical tests. The results revealed no statistically significant differences between the experimental groups with respect to age, gender, education, and survey completion device.

Questions and speeding feedback

We employed 15 closed questions on political solidarities and associated concepts that were adopted from scientific articles and established social surveys (see Goerres & Höhne, 2023). The 15 questions were thematically grouped: redistribution (two questions), governmental scope (five questions), social trust (three questions), and welfare chauvinism (five questions). The 15 questions were presented in the first quarter of the web survey and were accompanied by four-point, end-labeled rating scales with a vertical arrangement. Importantly, a randomly selected subset of respondents (about 50%) received the 15 closed questions with ascending rating scales and another randomly selected subset of respondents (about 50%) received the 15 closed questions with descending rating scales. We presented one question per web survey page (single question presentation). Appendix 1 includes the English translation of these questions and Appendix 2 displays screenshots of one of the questions on governmental scope including speeding feedback.

We kept the text of the speeding feedback identical across the three experimental groups and just varied the visual cues to not confound our treatment. In case of speeding, respondents were told that their answers appear quick and that careful reading of and careful thinking about the questions is important. In the feedback, we also asked respondents to allow for enough time. Appendix 1 includes the English translation of the speeding feedback.

We then specifically asked speeding respondents (i.e., respondents that received speeding feedback at least once) seven evaluative closed questions on how they perceived the feedback. These questions addressed the following seven aspects: 1) annoying, 2) confusing, 3) controlling, 4) attention-grabbing, 5) motivating, 6) lecturing, and 7) understanding. We presented all seven questions on one web survey page (item-by-item presentation) with five-point, end-labeled rating scales and a vertical arrangement. Appendix 1 includes the English translation of the evaluative closed questions.

At the end of the web survey, we asked all respondents four evaluative closed questions on how they perceived the web survey and its questions. These questions addressed the following four aspects: survey interest, survey difficulty, survey length, and topic sensitivity. We presented one question per web survey page (single question presentation) with seven-point, end-labeled rating scales and a vertical arrangement. Appendix 1 includes the English translation of the evaluative closed questions on the web survey.

Analytical strategy

In this study, we used STATA (version 18.0) for data preparation and analysis. We initially created a dichotomous variable for each of the 15 closed questions indicating whether respondents answered faster than the previously set speeding threshold: Speeding incident (1 = "Yes"). Similarly, we created dichotomous variables indicating whether respondents selected the first option of the scale: Primacy answer (1 = "Yes"). We then report descriptive statistics on speeding incidents, primacy answers, and speeding prompts received. This is done for each of the five question blocks and in total.

To examine our first research question, we investigate how speeding feedback with different visual cues affects speeding. Based on the 15 closed questions, we transformed the dataset into a long format, so that it consists of 15 observations (or answers) per respondent. We excluded answers to the first three questions from the analyses because respondents could receive the first speeding prompt only after the third question (or first question block). As answers (first level) are nested in respondents (second level), we used multilevel mixed-effects logistic regressions with random intercepts (see "melogit" command in STATA) and speeding incident (1 = "Yes") as dependent variable. We estimated four sequential models, stepwise adding independent variables. Model 1 is a null model and does not include any independent variables. We examine the Intraclass Correlation Coefficient (ICC) to determine the variation in speeding incident accounted for by respondent characteristics. Model 2 includes a variable indicating whether respondents received a speeding prompt in the previous question block (1 = "Yes"). In model 3, we additionally include the experimental groups in the form of neutral cue (1 = "Yes") and humanized cue (1 = "Yes") with no visual cue as reference. Finally, in model

4, we include a variable indicating whether respondents received the questions with a descending scale (1 = "Yes") and demographics in the form of female (1 = "Yes"), age (in years), medium education (1 = "Yes") and high education (1 = "Yes") with low education as reference. We also include the following self-reported survey evaluations as further independent variables: Survey interest (1 "Not interesting at all" to 7 "Very interesting"), difficulty (1 "Very easy" to 7 "Very difficult"), length (1 "Not long at all" to 7 "Very long"), and topic sensitivity (1 "Not sensitive at all" to 7 "Very sensitive"). These variables were included as previous research shows that they can affect respondents' answer behavior (Barth & Schmitz, 2021; Salvatore & Höhne, 2025; Zuell & Scholz, 2015). Appendix 3 reports the distribution of speeding incidents as well as poisson regressions with the number of speeding incidents as the dependent variable (only considering respondents who received at least one speeding prompt).

To investigate our second research question, we examine how speeding feedback with different visual cues affects data quality in terms of primacy effects. For this purpose, we again used multilevel mixed-effects logistic regressions with random intercepts and primacy answer (1 = "Yes") as dependent variable. We followed the same modelling approach and analytical strategy as in the analysis on our first research question. Appendix 3 reports the distribution of primacy answers as well as poisson regressions with the number of primacy answers as the dependent variable (only considering respondents who received at least one speeding prompt).

To examine our third research question, we investigate how speeding feedback with different visual cues is evaluated by respondents. Specifically, we investigate the extent to which respondents perceived the speeding feedback as annoying, confusing, controlling, attention-grabbing, motivating, lecturing, and understanding by experimental group. These evaluations were employed as closed questions with five-point rating scales (1 "Applies not at all" to 5 "Applies strongly"). For the analysis, we only considered respondents who received at least one speeding prompt (n = 1,559). Based on the seven evaluative closed questions, we conducted one-way analyses of variance (ANOVAs) using the Bonferroni α-inflation correction procedure.

In our analysis, we use a p-level smaller than 0.05 to determine statistical significance, except for the multilevel regressions in which we use a p-value smaller than 0.01. The reason is the relatively high number of observations (about 24,000) increasing statistical power (see Miller & Ulrich, 2019).

Results

Descriptive statistics

In a first step, we examine descriptive statistics on speeding incidents, primacy answers, and speeding prompts received across the three experimental groups. Table 1 shows these statistics for each of the five question blocks and in total. On average, respondents answered three questions faster than the speeding threshold, provided three primacy answers, and received two speeding prompts. Importantly, the average number of speeding incidents varies substantially between the five question blocks, ranging from 0.06 (fifth block) to 0.91 (first block). This also applies to the percentage of respondents receiving a speeding prompt. While about 50% of respondents received a speeding prompt in the first, third, and fourth question block, respectively, only between 5% and 16% of respondents received a speeding prompt in the second and fifth question block. The average number of primacy answers, in contrast, varies between 0.47 (third block) and 0.72 (second block). Importantly, there are almost no differences between the three experimental groups when it comes to speeding incidents, primacy answers, and speeding prompts received.

Research question 1

Regarding our first research question, we investigated how speeding feedback with different visual cues affects speeding. To this end, we estimated two-level mixed-effects logistic regressions with random intercepts and speeding incident (1 = "Yes") as dependent variable. Table 2 displays the results of our four sequential models.

In model 1 (null model), the ICC indicates that a substantial proportion of the variation in speeding incident is accounted for by the respondent level (ICC = 0.31). Looking at the second model, receiving a speeding prompt results in a lower likelihood of speeding incident. In line with previous research, this indicates that speeding prompts indeed reduce speeding. The experimental groups, as indicated by the third model, are not associated with speeding incident, suggesting that visual cues do not affect speeding. This is in line with our descriptive statistics showing that the number of speeding incidents and speeding prompts received do not vary across the three experimental groups. In the fourth model, we now include independent variables on scale direction, demographics, and self-reported survey evaluations. Descending scale and high education are both positively associated with speeding incident, while age is negatively associated with speeding incident. Survey interest and difficulty are also negatively associated with speeding incident, suggesting that both respondents evaluating the survey as interesting and respondents evaluating the survey as difficult are less prone to speeding.

Research question 2

In the context of our second research question, we investigate how speeding feedback with different visual cues affects data quality in terms of primacy effects. Similar to the analyses on our first research question, we estimated two-level mixed-effects logistic regressions with random intercepts and primacy answer (1 = "Yes") as dependent variable. The results of the four sequential models are presented in Table 3.

The ICC of model 1 (null model) indicates that a substantial proportion of the variation in primacy answer is accounted for by the respondent level (ICC = 0.27). Looking at the second model, receiving a speeding prompt is positively associated with primacy answer. This indicates that speeding prompts may negatively affect data quality. In the third model, as in our previous analyses on speeding incidents, experimental groups varying with respect to the visual cues are not associated with primacy answer. In the fourth model, we additionally include independent variables on scale direction, demographics, and self-reported survey evaluations. However, only descending scale is (positively) associated with primacy answer, suggesting that this particular scale direction obtains more primacy answers than its ascending counterpart.

Research question 3

Finally, we investigate how speeding feedback with different visual cues is evaluated by respondents. Table 4 shows the results of one-way analyses of variance (ANOVAs), including the mean evaluations across experimental groups and pairwise differences between groups. On

Table 1. Descriptive statistics on speeding incidents, primacy answers, and speeding prompts received

	No visual cue	Neutral cue	Humanized cue	Total
Across all five question blocks				
Average number of speeding incidents	2.76	2.59	2.95	2.77
Average number of primacy answers	3.16	2.94	3.05	3.05
Average number of speeding prompts	1.72	1.71	1.86	1.76
1st question block				
Average number of speeding incidents	0.87	0.85	1.01	0.91
Average number of primacy answers	0.70	0.65	0.68	0.67
Percentage of speeding prompts	52	53	58	54
2nd question block				
Average number of speeding incidents	0.25	0.21	0.28	0.25
Average number of primacy answers	0.70	0.75	0.72	0.72
Percentage of speeding prompts	16	15	19	16
3rd question block				
Average number of speeding incidents	0.87	0.80	0.87	0.85
Average number of primacy answers	0.49	0.44	0.47	0.47
Percentage of speeding prompts	52	51	55	53
4th question block				
Average number of speeding incidents	0.70	0.67	0.72	0.70
Average number of primacy answers	0.75	0.67	0.67	0.70
Percentage of speeding prompts	46	46	49	47

Table 1 (continued).

Evaluation	No visual cue	Neutral cue	Humanized cue	Total
5th question block				
Average number of speeding incidents	0.07	0.06	0.07	0.06
Average number of primacy answers	0.52	0.43	0.51	0.48
Percentage of speeding prompts	6	5	6	5

Table 2. Two-level mixed-effects logistic regressions with random intercepts and speeding incident (1 = "Yes") as the dependent variable

	M1		M2		M3		M4	
	Logit	SE	Logit	SE	Logit	SE	Logit	SE
Intercept	-2.14**	0.04	-1.94**	0.05	-1.94**	0.08	1.66**	0.29
Prompt in previous question block (reference: no prompt)			-0.80**	0.06	-0.80**	0.06	-0.83**	0.06
Experimental group (reference: no visual cue)								
Neutral cue					-0.10	0.10	-0.14	0.09
Humanized cue					0.10	0.10	0.00	0.09
Descending scale (reference: ascending scale)							0.22*	0.07
Female (reference: male)							0.19*	0.07
Age (in years)							-0.05**	0.00
Education (reference: low education)								
Medium education							-0.02	0.10
High education							0.36**	0.10
Survey interest (1 "Not interesting at all" to 7 "Very interesting")							-0.08*	0.03
Survey difficulty (1 "Very easy" to 7 "Very difficult")							-0.20**	0.03
Survey length (1 "Not long at all" to 7 "Very long")							0.02	0.02
Topic sensitivity (1 "Not sensitive at all" to 7 "Very sensitive")							-0.04	0.02
Respondent-level ICC	0.31		0.44		0.44		0.31	
Observations	23,928		23,928		23,928		23,928	

Note. **p < 0.001, *p < 0.01. SE = Standard error. Exclusion of respondents with missing values for any independent variable.

Table 3. Two-level mixed-effects logistic regressions with random intercepts and primacy answer (1 = "Yes") as the dependent variable

	M1		M2		M3		M4	
	Logit	SE	Logit	SE	Logit	SE	Logit	SE
Intercept	-1.72**	0.03	-1.78**	0.04	-1.74**	0.06	-1.99**	0.25
Prompt in previous question block (reference: no prompt)			0.15*	0.04	0.15*	0.04	0.16**	0.04
Experimental group (reference: no visual cue)								
Neutral cue					-0.08	0.08	-0.09	0.08
Humanized cue					-0.04	0.08	-0.03	0.08
Descending scale (reference: ascending scale)							0.33**	0.06
Female (reference: male)							-0.05	0.06
Age (in years)							0.00	0.00
Education (reference: low education)								
Medium education							-0.08	0.08
High education							-0.19	0.09
Survey interest (1 "Not interesting at all" to 7 "Very interesting")							0.06	0.03
Survey difficulty (1 "Very easy" to 7 "Very difficult")							-0.05	0.03
Survey length (1 "Not long at all" to 7 "Very long")							0.00	0.02
Topic sensitivity (1 "Not sensitive at all" to 7 "Very sensitive")							-0.01	0.02
Respondent-level ICC	0.27		0.28		0.27		0.27	
Observations	23,887		23,887		23,887		23,887	

Note. **p < 0.001, *p < 0.01. SE = Standard error. Exclusion of respondents with missing values for any independent variable.

average, the aspects annoying ($\bar{x}$ = 3.59) and lecturing ($\bar{x}$ = 3.56) received the highest ratings, followed by controlling ($\bar{x}$ = 3.25) and having understanding ($\bar{x}$ = 3.08). In contrast, the aspects motivating ($\bar{x}$ = 1.90), confusing ($\bar{x}$ = 2.36), and attention-grabbing ($\bar{x}$ = 2.37) received the lowest ratings. However, we do not observe any differences between the three experimental groups, except for attention-grabbing. Respondents receiving speeding feedback with a neutral cue rated the feedback as more attention-grabbing than respondents receiving speeding feedback with a humanized cue. Overall, respondents do not seem to evaluate the three speeding feedback designs differently.

Discussion and conclusion

The aim of this study was to investigate speeding feedback and its efficacy in terms of decreasing speeding and increasing data quality. In addition, we measured respondents' disposition towards or perception of speeding feedback. To this end, we focused on three research questions and used an experimental design in which we randomly assigned respondents to one of three speeding design groups. Our results show that speeding feedback reduces speeding, regardless of the visual cue. However, speeding feedback does not improve data quality and is partly evaluated negatively by respondents.

With respect to our first research question on how speeding feedback with different visual cues (i.e., no cue, neutral cue, or humanized cue) affects speeding, we found that speeding feedback works by slowing respondents down when answering subsequent questions. However, in contrast to previous research (Zhang & Canrad, 2018), we did not find differences across the visual cues that we employed. To put it differently, the inclusion of visual cues in the form of a paper clip (neutral cue) or a photo of the principal investigator (humanized cue) did not further reduce speeding (or increase the efficacy of speeding feedback). While the design of the speeding feedback had no impact, scale direction (i.e., descending or ascending) had an impact on speeding. More specifically, descending scales resulted in more speeding than their ascending counterparts. One explanation is that the question stems had a positive or balanced formulation and that the descending scales started with a positively formulated answer option (e.g., "Most people try to behave fair"). Accordingly, it may have taken respondents less time to select the first answer option that seems reasonable resulting in more speeding incidents. We also found that higher educated respondents (with potentially higher cognitive skills) are more inclined to speed, whereas older respondents (with potentially lower cognitive skills) are less inclined to speed. Respondents evaluating web survey participation as more interesting and difficult are also less inclined to speed. Overall, these findings correspond to the survey satisficing framework (Krosnick, 1991) suggesting that respondents' ability (or cognitive skills), motivation (or survey interest), and task (or survey) difficulty matter when it comes to speeding.

Our second research question on how speeding feedback with different visual cues affects data quality focused on primacy effects (i.e., selecting the first answer option of the scale). For this purpose, we additionally randomly assigned respondents to different scale directions (i.e., ascending or descending). Our findings showed that speeding feedback did not decrease primacy answers to later questions. Interestingly, we also found that descending scales were more prone to primacy answers than their ascending counterparts. This finding is in line with findings reported by Krebs and Hoffmeyer-Zlotnik (2010) as well as Krebs and Höhne (2020)

Table 4. Mean differences of speeding feedback evaluations between the three experimental groups

Evaluation	No visual cue (1)	Neutral cue (2)	Humanized cue (3)	F value ($df_1 = 2$)	df_2	Pairwise differences		
						(1) vs. (2)	(1) vs. (3)	(2) vs. (3)
Annoying	3.62	3.48	3.67	2.74	1554	0.14	-0.05	-0.19
Confusing	2.36	2.40	2.34	0.21	1544	-0.04	0.02	0.06
Controlling	3.25	3.14	3.37	2.72	1549	0.11	-0.12	-0.23
Attention-grabbing	2.36	2.48	2.26	3.33*	1548	-0.12	0.10	0.22*
Motivating	1.94	1.94	1.83	1.42	1548	0.00	0.11	0.11
Lecturing	3.47	3.55	3.64	1.66	1550	-0.08	-0.16	-0.09
Understanding	3.04	3.14	3.06	0.67	1547	-0.10	-0.01	0.08

Note. *$p < 0.05$. One-way analyses of variance (ANOVA) using the Bonferroni α-inflation correction procedure. Only including respondents who received at least one speeding prompt ($n = 1,559$). Feedback evaluations were measured using five-point rating scales: 1 “Applies not at all” to 5 “Applies strongly.”

showing that primacy effects are more common for descending than ascending scales. One possible explanation is that these scales additionally foster a kind of "positivity bias" (Tourangeau, Rips, & Rasinski, 2000), as they start with the positively formulated answer option. Regardless of the effect to which respondents' answer behavior in descending scales is attributable, it compromises data quality. We therefore argue for the use of ascending scales instead.

Regarding our third research question on how speeding feedback with different visual cues is evaluated by respondents, we investigated respondents' perceptions of and attitudes towards the following seven aspects by experimental group: annoying, confusing, controlling, attention-grabbing, motivating, lecturing, and understanding. Respondents provided the highest ratings for annoying, lecturing, and controlling, indicating that they do not appreciate speeding feedback. In addition, speeding feedback did not seem to motivate respondents, as the motivating aspect received the lowest rating of all seven aspects under investigation. However, it appears that respondents somehow showed understanding for feedback, as the understanding aspect received the fourth highest rating. Even though this study sheds new light on how speeding feedback is evaluated by respondents, it remains unclear how these evaluations are related to respondents' answer behavior. We therefore suggest that future studies relate respondents' evaluations to their speeding and answer behavior. In addition, we recommend including follow-up probes in the form of open narrative questions to gather more nuanced information on the perception of speeding feedback in web surveys.

Our study has some methodological limitations that provide new avenues for future research. First, in this study, we set one single speeding threshold in the form of 300 milliseconds (msec) per word that was adopted from Zhang and Conrad (2018). As shown in our results section, this resulted in some variation regarding speeding incidents and speeding prompts received across the five question blocks. For example, the first question block included the highest number of speeding incidents and speeding prompts, whereas the fifth question block included the lowest number. The reason is that the questions in these blocks differ in word length, which in turn affects the speeding thresholds. We therefore suggest further investigating thresholds that are not exclusively determined by the number of words but also take question difficulty into consideration. Second, we grouped our 15 closed questions into five question blocks (with three questions each) and prompted respondents at the end of the block in case of speeding incidents. From our perspective, it would be worthwhile investigating somewhat less intrusive schemes, such as only providing feedback on the first speeding occasion or other alternated schemes. This potentially leads to more positive evaluations by respondents, while still effectively reducing speeding. Relatedly, it would be also possible to further "humanize" speeding feedback by building on the media capabilities of contemporary electronic devices. To put it differently, it would be also possible to provide a pre-recorded video (of the PI) briefly outlining the importance of thoughtful answering. Third, we exclusively focused on data quality in terms of primacy effects. In future studies, it would be good to consider further data quality indicators. Specifically, researchers may consider further types of survey satisficing, such as non-differentiation and "Don't know" answers (Krosnick, 1991), or investigate the relation between speeding and reliability and validity. For example, the inspection of correlations between target and criterion questions can shed light on criterion validity (Höhne & Yan, 2020; Yeager & Krosnick, 2012).

Our findings provided new evidence on the efficacy of speeding feedback with different visual cues. The provision of feedback effectively slows respondents down in answering later questions so that it has a lasting impact. The visual cues did not affect speeding. We therefore conclude that plain, text-based feedback without visual cues is preferable. Considering the fact that speeding feedback did not improve data quality and was not favored by respondents, we urge survey researchers and practitioners to critically reflect upon its implementation. In doing so, it would be wise to decide on a question-level and to, for example, conduct pretests that inform about appropriate speeding thresholds and the efficacy of speeding feedback. In addition, we encourage other researchers to further conduct methodological research on speeding feedback in self-administered web surveys, as it has proven its worth in other studies (Conrad et al., 2017; Zhang & Conrad, 2018).

Author contributions

Original conception of the overall study: JKH and AG; Experimental and survey design: JKH and AG; Survey programming and pretesting: JC and JKH; Statistical analysis: JC and JKH; Article writing: JKH, AG and JC.

References

Barth, A., & Schmitz, A. (2021). Interviewers' and respondents' joint production of response quality in open-ended questions: A multilevel negative-binomial regression approach. Methods, Data, Analysis, 15, 43-76. https://doi.org/10.12758/mda.2020.08

Conrad, F. G., Tourangeau, R., Couper, M. P., & Zhang, C. (2017). Reducing speeding in web surveys by providing immediate feedback. Survey Research Methods, 11(1), 45–61. https://doi.org/10.18148/srm/2017.v11i1.6304

Couper, M. P. (2000). Usability evaluation of computer-assisted survey instruments. Social Science Computer Review, 18(4), 384–396. https://doi.org/10.1177%2F089443930001800402

Goerres, A., & Höhne, J.K. (2023). Evaluating the response effort and data quality of established political solidarity measures: A pre-registered experimental test in an online survey of the German adult resident population in 2021. Quality and Quantity. https://doi.org/10.1007/s11135-022-01594-4

Höhne, J. K., Schlosser, S., Couper, M. P., & Blom, A. G. (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

Höhne, J.K., & Yan, T. (2020). Investigating the impact of violations of the "left and top means first" heuristic on response behavior and data quality. International Journal of Social Research Methodology, 23, 347–353. https://doi.org/10.1080/13645579.2019.1696087

Krebs, D., & Hoffmeyer-Zlotnik, J. H. P. (2010). Positive first or negative first? Effects of the order of answering categories on response behavior. Methodology, 6, 118–127. https://doi.org/10.1027/1614-2241/a000013

Krebs, D., & Höhne, J. K. (2020). Antwortskalenrichtung und Umfragemodus. In A. Mays, A. Dingelstedt, V. Hambauer, S. Schlosser, F. Berens, J. Leibold, & J. K. Höhne (Eds.). Grundlagen—Methoden—Anwendungen in den Sozialwissenschaften (pp. 231–246). Springer VS. https://doi.org/10.1007/978-3-658-15629-9_12

Krosnick, J. A., (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology, 5(3), 213–236. https://doi.org/10.1002/acp.2350050305

Malhotra, N. (2009). Completion Time and Response Order Effects in Web Surveys. Public Opinion Quarterly, 72(5), 914–934. https://doi.org/10.1093/poq/nfn050

Miller, J., & Ulrich, R (2019). The quest for an optimal alpha. PLoS ONE 14, e0208631. https://doi.org/10.1371/journal.pone.0208631

Nass, C., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. Journal of Applied Social Psychology, 29, 1093–1110. https://doi.org/10.1111/j.1559-1816.1999.tb00142.x

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. Journal of Applied Social Psychology, 27, 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Reeves, B., & Nass, C. (1997). The media equation: how people treat computers, television, and new media like real people and places. Cambridge: CSLI and Cambridge University Press.

Saarijärvi, M., & Bratt, E.-L. (2021). When face-to-face interviews are not possible: Tips and tricks for video, telephone, online chat, and email interviews in qualitative research. European Journal of Cardiovascular Nursing, 20, 392-396. https://doi.org/10.1093/eurjcn/zvab038

Salvatore, C., & Höhne, J.K. (2025). Explaining item-nonresponse in open questions with requests for voice responses. In Pollice, A., & Mariani, P. (eds.), Methodological and Applied Statistics and Demography IV. Cham: Springer. https://doi.org/0.1007/978-3-031-64447-4_82

Schober, M. F. (2018). The future of face-to-face interviewing. Quality Assurance in Education, 26(2), 290-302. https://doi.org/10.1108/QAE-06-2017-0033

Schlosser, S., & Höhne, J. K. (2018). ECSP – Embedded Client Side Paradata. Zenodo. https://doi.org/10.5281/zenodo.1218941

Schlosser, S., & Höhne, J.K. (2020). ECSP – Embedded Client Side Paradata. Zenodo. https://doi.org/10.5281/zenodo.3782592

Self, B. (2021). Conducting interviews during the Covid-19 pandemic and beyond. Forum Qualitative Social Research, 22(3), Article 15. https://doi.org/10.17169/fqs-22.3.3741

Sendelbah, A., Vehovar, V., Slavec, A., & Petrovčič, A. (2016). Investigating respondent multitasking in web surveys using paradata. Computers in Human Behavior, 55, 777–787. https://doi.org/10.1016/j.chb.2015.10.028

Shamon, H., & Berning, C. C. (2020). Attention Check Items and Instructions in Online Surveys with Incentivized and Non-Incentivized Samples: Boon or Bane for Data Quality?. Survey Research Methods, 14(1), 55–77. https://doi.org/10.18148/srm/2020.v14i1.7374

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). Thinking about answers: The application of cognitive processes to survey methodology. San Francisco: Jossey-Bass Publishers.

Tourangeau, R., Couper, M. P., & Steiger, D. M. (2003). Humanizing self-administered surveys: Experiments on social presence in web and IVR surveys. Computers in Human Behavior, 19, 1–24. https://doi.org/10.1016/S0747-5632(02)00032-8

Tourangeau, R., Rips, L. L., and Rasinski, K. (2000). The Psychology of Survey Response. New York, NY: Cambridge University Press.
Yeager, D. S., & Krosnick, J. A. (2012). Does mentioning "some people" and "other people" in an opinion question improve measurement quality? Public Opinion Quarterly, 76, 131–141. https://doi.org/10.1093/poq/nfr066
Zhang, C., & Conrad, F. G. (2014). Speeding in web surveys: The tendency to answer very fast and its association with straightlining. Survey Research Methods, 8, 127–135. https://doi.org/10.18148/srm/2014.v8i2.5453
Zhang, C., & Conrad, F. G. (2018). Intervening to reduce satisficing behaviors in web surveys: Evidence from two experiments on how it works. Social Science Computer Review, 36(1), 57–81. https://doi.org/10.1177/0894439316683923
Zuell, C., & Scholz, E. (2015). Who is willing to answer open-ended questions on the meaning of left and right? Bulletin de Méthodologie Sociologique, 127, 26-42. https://doi.org/10.1177/0759106315582199

Appendix 1

English translations of the closed questions, speeding feedback, and evaluative questions.

Closed questions for which speeding feedback was provided (ascending scale only)

1) Now please indicate to what extent the following things should be the responsibility of the state. The state should reduce the income gap between rich and poor. (redistribution 1)
   Rating scale: 1 "Not being responsible" to 4 "Being responsible"
2) Here are two statements about a controversial issue and a scale that you can use to grade your own opinion about it. If you completely agree with the statement above the scale, select the answer box at the top. If you completely agree with the statement below the scale, select the answer box at the bottom. If your opinion is somewhere in between, you can express this with one of the answer boxes in between. (redistribution 2)
   Rating scale: 1 "The state should not take more responsibility for ensuring that every citizen is covered" to 4 "The state should take more responsibility for ensuring that every citizen is covered"
3) People have different ideas about what the state should and should not be responsible for. For each of the following tasks, please tell us whether the state should be responsible for it. Should the state be responsible for ensuring a decent standard of living in old age? (governmental scope 1)
   Rating scale: 1 "Not being responsible" to 4 "Being responsible"
4) Should the state be responsible for ensuring a decent standard of living in young age? (governmental scope 2)
   Rating scale: 1 "Not being responsible" to 4 "Being responsible"
5) Should the state be responsible for ensuring childcare options for working parents? (governmental scope 3)
   Rating scale: 1 "Not being responsible" to 4 "Being responsible"

6) Should the state be responsible for ensuring a decent standard of living for the unemployed? (governmental scope 4)
Rating scale: 1 "Not being responsible" to 4 "Being responsible"
7) Should the state be responsible for ensuring a decent standard of living for poor people? (governmental scope 5)
Rating scale: 1 "Not being responsible" to 4 "Being responsible"
8) In general, do you think that most people can be trusted, or that you can't be careful enough when dealing with other people? (social trust 1)
Rating scale: 1 "You can't be too careful" to 4 "Most people can be trusted"
9) Do you think most people try to take advantage of you when they have the opportunity, or do most people try to be fair? (social trust 2)
Rating scale: 1 "Most people try to take advantage of me" to 4 "Most people try to behave fair"
10) Do you think that people mostly try to be helpful, or that people mostly look out for their own advantage? (social trust 3)
Rating scale: 1 "People are mostly looking out for their own advantage" to 4 "People mostly try to be helpful"
11) Immigrants from outside the EU should have the same entitlement to social welfare in the future as people born in Germany. (welfare chauvinism 1)
Rating scale: 1 "Disagree" to 4 "Agree"
12) Immigrants from the EU should have the same entitlement to social welfare in the future as people born in Germany. (welfare chauvinism 2)
Rating scale: 1 "Disagree" to 4 "Agree"
13) Despite the welfare state, people are hard-working. (welfare chauvinism 3)
Rating scale: 1 "Disagree" to 4 "Agree"
14) Despite the welfare state, people look after themselves.
Rating scale: 1 "Disagree" to 4 "Agree"

The state should increase social benefits. (welfare chauvinism 5)
Rating scale: 1 "Disagree" to 4 "Agree"

Speeding feedback
Your answers sometimes seem very quick. It is important that you read each question carefully and think carefully about your answers. Please allow enough time for all questions.

Depending on the experimental group, the speeding feedback was provided without a visual cue (control), a neutral cue (paper clip), or a humanized cue (photo of PI).

Evaluative closed questions on the speeding feedback perception
While answering the last 15 questions, you were asked to read the questions carefully and to think about your answers carefully. Please tell us how you felt about these prompts.

1) I find such prompts annoying.
2) I find such prompts confusing.
3) I feel controlled by such prompts.
4) I increase my attention with such prompts.

5) I find such prompts motivating.
6) I feel lectured by such prompts.
7) I have understanding for such prompts.
   Rating scale: 1 "Does not apply" to 5 "Applies"

Evaluative closed questions on the web survey perception

1) How interesting did you find the web survey?
   Rating scale: 1 "Very interesting" to 7 "Not at all interesting"
2) How easy or difficult did you find it to answer the questions asked?
   Rating scale: 1 "Very easy" to 7 "Very difficult"
3) How long did you find the web survey?
   Rating scale: 1 "Very long" to 7 "Not at all long"
4) How personal did you find it to answer the questions asked?
   Rating scale: 1 "Very personal" to 7 "Not at all personal"

Note. The order of the questions in the web survey corresponds to the presentation order in Appendix 1. The 15 closed questions for which we provided speeding feedback were presented in a single question presentation format (one question per web survey page) and the seven evaluative questions were presented in an item-by-item presentation format (all questions on the same web survey page). All questions were presented with vertically aligned rating scales. The original German question wordings are available from the second author upon request.

Appendix 2

Exemplary screenshots of the closed question on governmental scope including speeding feedback.

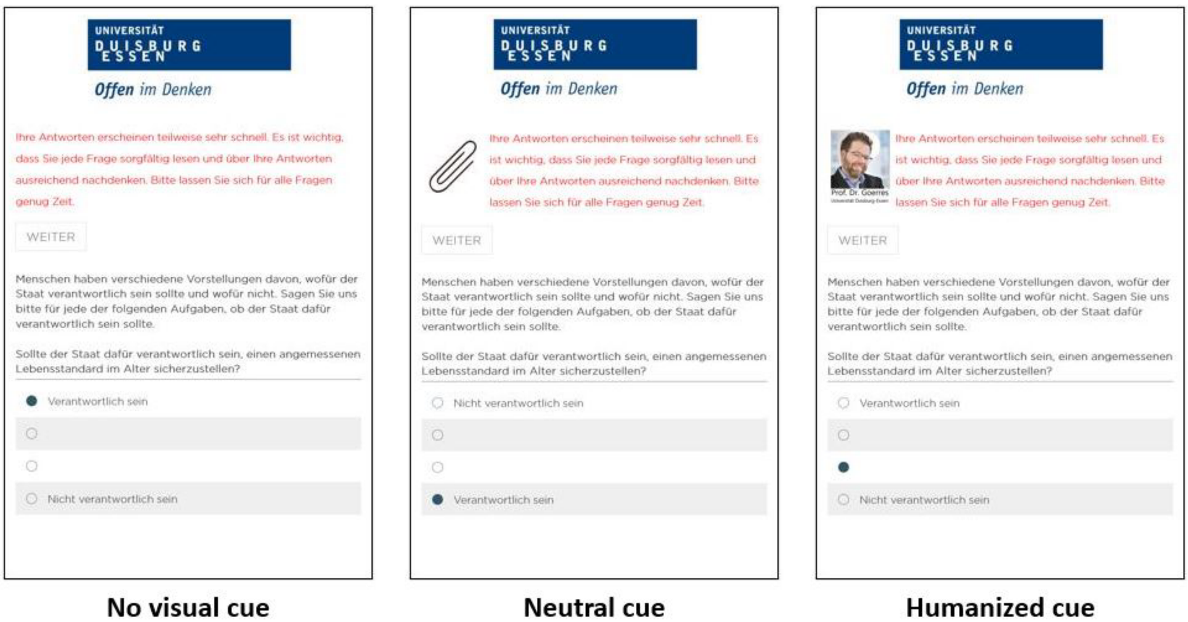

Figure A1. Example screenshots of the closed question on governmental scope across the three experimental groups

Note. No visual cue (control) on the left, neutral cue (paper clip) in the middle, and humanized cue (photo of PI) on the right. Screenshots were taken on a smartphone.

Appendix 3

Analyses considering only respondents who received at least one speeding prompt

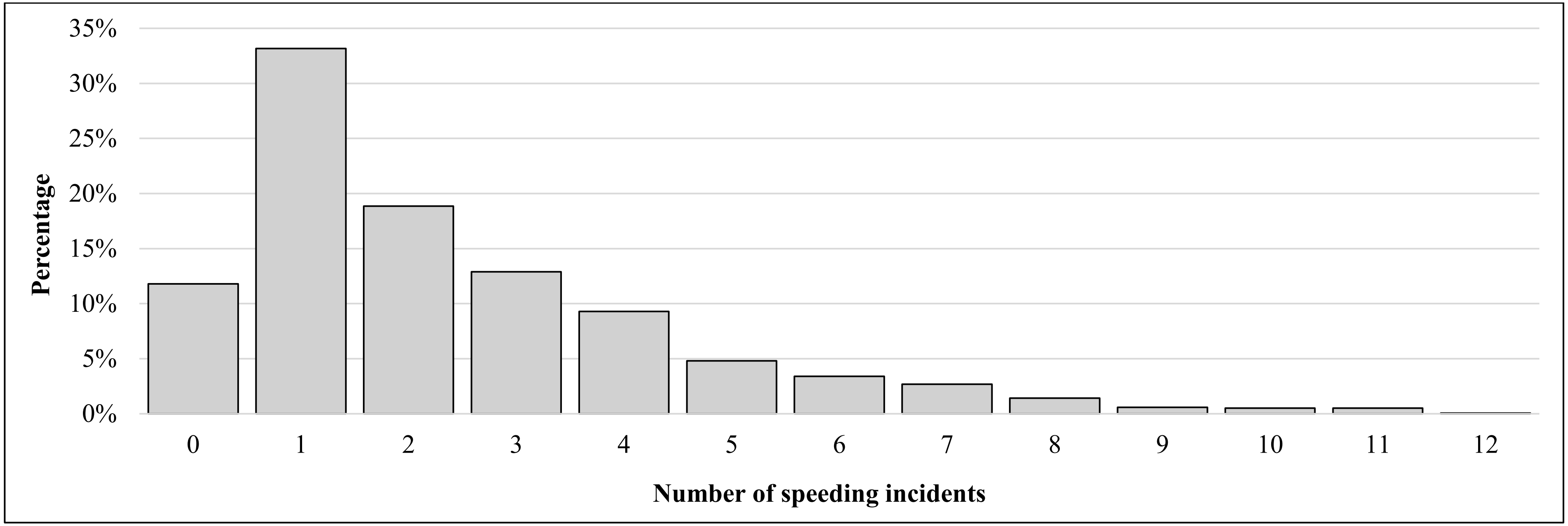

Figure A2. Distribution of the number of speeding incidents in percentages

Note. We only considered respondents who received at least one speeding prompt (n = 1,559). In addition, we excluded speeding incidents related to the first three questions because respondents could receive the first speeding prompt only after the third question (or first question block). Therefore, “zero” speeding incidents indicate that respondents received a speeding prompt after the first question block but did not have any other speeding incidents after the first question block.

Table A1. Poisson regression with number of speeding incidents (0-12) as the dependent variable

	M1		M2	
	Coefficient	SE	Coefficient	SE
Intercept	0.90**	0.03	2.42**	0.13
Experimental group (reference: no visual cue)				
Neutral cue	-0.08	0.04	-0.09*	0.04
Humanized cue	-0.03	0.04	-0.05	0.04
Descending scale (reference: ascending scale)			0.07*	0.03
Female (reference: male)			0.08*	0.03
Age (in years)			-0.02**	0.00
Education (reference: low education)				
Medium education			-0.10*	0.05
High education			0.07	0.05
Survey interest (1 “Not interesting at all” to 7 “Very interesting”)			-0.04**	0.01
Survey difficulty (1 “Very easy” to 7 “Very difficult”)			-0.08**	0.01
Survey length (1 “Not long at all” to 7 “Very long”)			0.00	0.01
Topic sensitivity (1 “Not sensitive at all” to 7 “Very sensitive”)			-0.01	0.01
Pseudo R^2	0.00		0.10	
Observations	1,548		1,548	

Note. **p < 0.01, *p < 0.05. SE = Standard error. Exclusion of respondents with missing values for any independent variable.

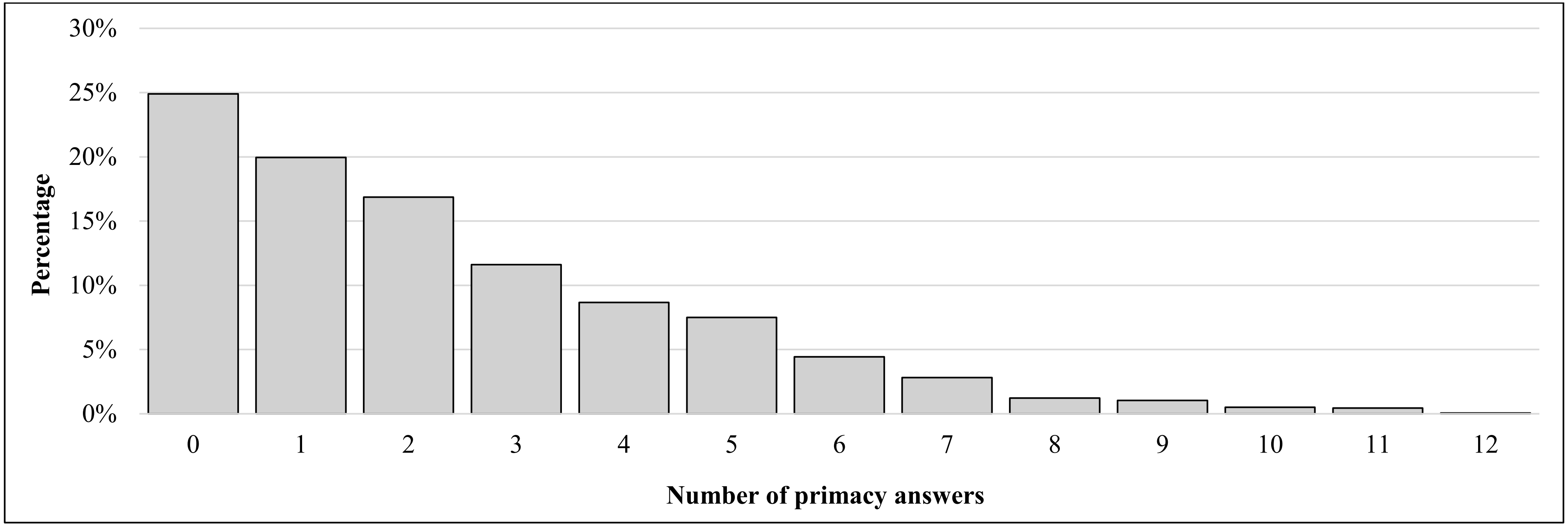

Figure A3. Distribution of the number of primacy answers in percentages

Note. We only considered respondents who received at least one speeding prompt (n = 1,559). In addition, we excluded primacy answers related to the first three questions because respondents could receive the first speeding prompt only after the third question (or first question block).

Table A2. Poisson regression with number of primacy answers (0-12) as the dependent variable

	M1		M2	
	Coefficient	SE	Coefficient	SE
Intercept	0.88**	0.03	0.84**	0.13
Experimental group (reference: no visual cue)				
Neutral cue			-0.04	0.04
Humanized cue			-0.00	0.04
Descending scale (reference: ascending scale)			0.22**	0.03
Female (reference: male)			-0.02	0.03
Age (in years)			-0.00	0.00
Education (reference: low education)				
Medium education			-0.04	0.05
High education			-0.08	0.05
Survey interest (1 "Not interesting at all" to 7 "Very interesting")			0.05**	0.01
Survey difficulty (1 "Very easy" to 7 "Very difficult")			-0.05**	0.01
Survey length (1 "Not long at all" to 7 "Very long")			-0.00	0.01
Topic sensitivity (1 "Not sensitive at all" to 7 "Very sensitive")			-0.01	0.01
Pseudo R^2	0.00		0.01	
Observations	1,548		1,548	

Note. **p < 0.01, *p < 0.05. SE = Standard error. Exclusion of respondents with missing values for any independent variable.