Journal of Survey Statistics and Methodology (2017) 5, 48–69

RESPONSE BEHAVIOR IN A VIDEO-WEB SURVEY: A MODE COMPARISON STUDY

MARIEKE HAAN*
YFKE P. ONGENA
JORRE T. A. VANNIEUWENHUYZE
KEES DE GLOPPER

Intrigued by the possibilities of improving the data quality of web surveys by incorporating human-like features, we developed a video-web survey for this study. This paper describes an experiment that compares response behavior in the video-web mode to traditional web and interviewer-administered surveys. The disclosure of sensitive information and respondents' engagement were examined. Overall, despite the visual and auditory representation of a human interviewer in the video-web mode, video-web seems to have been experienced by respondents much like a traditional web survey. Based on these results, we argue that for human-like features to fully increase the level of engagement it would require the inclusion of responsiveness. However, researchers should be aware of possible social presence effects that may arise when creating web surveys with responsive human-like features.

KEYWORDS: Mode comparison; Satisficing; Social desirability; Video-web survey.

1. INTRODUCTION

Web surveys are a popular method for data collection in survey research. Survey organizations are increasingly using web surveys because of the time and cost savings (Callegaro, Baker, Bethlehem, Goritz, Krosnick, et al. 2014),

Marieke Haan and Jorre Vannieuwenhuyze are with the Department of Methodology and Statistics, Utrecht University, the Netherlands. Yfke Ongena and Kees de Glopper are with the Department of Communication Sciences, University of Groningen, the Netherlands. This research is part of a project that was funded by the Netherlands Organization for Scientific Research (NWO), Grant #471-09-002. The authors want to thank the editor of JSSAM, Roger Tourangeau, and the reviewers for their helpful suggestions and their constructive feedback.
*Address correspondence to Marieke Haan, Department of Methodology and Statistics, Utrecht University, the Netherlands; E-mail: m.haan@uu.nl.

doi: 10.1093/jssam/smw023 Advance access publication 28 December 2016

the ever-improving technological possibilities, and the increase in Internet use in Western countries. In addition, in self-administered questionnaires (SAQs) such as web questionnaires, respondents are inclined to answer especially sensitive questions more truthfully than in interviewer-administered modes (e.g., Tourangeau and Smith 1996). However, findings on the risks for satisficing (i.e., not optimally responding to questions, see Krosnick 1991) across survey modes are mixed. On the one hand, the likelihood for satisficing may be greater in a web survey than in interviewer-administered modes because the motivation to participate in a web survey is often lower, as is shown by the low response rates for web surveys (Lozar Manfreda, Bosnjak, Berzelak, Haas, and Vehovar 2008). Moreover, when respondents do participate, they can easily be distracted from the task, and there is no interviewer present who has control over the interview situation (Heerwegh and Loosveldt 2008). Consequently, respondents can essentially act at their own discretion without having to justify their answers to an interviewer. On the other hand, satisficing may be less prevalent in web than in interviewer-administered modes because web surveys allow respondents to complete the questionnaire at their own chosen time and at their own pace (Holbrook, Green, and Krosnick 2003; Fricker, Galesic, Tourangeau, and Yan 2005).

Social interface theory states that the introduction of humanizing cues in a computer interface can change the reactions of users of that interface; their responses become more similar to responses in human-to-human interaction (Sproull, Subramani, Kiesler, Walker, and Waters 1996; Reeves and Nass 1997; Tourangeau, Couper, and Steiger 2003). Based on this theory, researchers have integrated features that simulate attributes or behaviors of humans in web surveys. Previous work suggests that these human-like features may improve the level of engagement respondents experience when participating in a web survey, while still maintaining the feeling of being able to answer the questions anonymously (Tourangeau et al. 2003; Tourangeau, Conrad, and Couper 2013).

Human-like features in web surveys can take various forms. First, they can be static, which means there is no movement or change in what is displayed to the respondent (e.g., a picture of an interviewer), or they can be dynamic, which entails some kind of movement or change (e.g., a virtual interviewer). Second, human-like cues may contain auditory stimuli (e.g., a human voice) and/or visual stimuli (e.g., a human face). Third, these human-like features can differ in their responsiveness, the degree to which they respond to the respondent's action (Tourangeau et al. 2013). Nonresponsive human-like cues merely confer human characteristics such as gender, race, and social class to the interface, while responsive human-like cues give respondents the impression that the interface is an agent that actively interacts with the respondent. A nonresponsive cue may, based on the mere presence hypothesis, induce a priming effect (priming specific attitudes; see Tourangeau et al. 2003), whereas responsive cues are more obviously humanizing and are more likely to induce social presence effects—for example,

effects based on the notion that "people avoid expressing negative stereotypes in the presence of a member of the target group—the interviewer—for fear of giving offense" (Tourangeau et al. 2003, p. 11).

One way of humanizing web surveys is by including pictures of the researchers and interviewers; this is a static nonresponsive feature with visual human aspects. Including pictures in web surveys does not seem to yield differences in sensitive reporting compared with traditional web surveys (Krysan and Couper 2006; Tourangeau et al. 2003). Second, web surveys can be humanized by adding a human voice, a dynamic but nonresponsive feature with auditory human features. Since the 1990s, researchers have been using audio computer self-interviewing modes (ACASI), in which respondents listen to a prerecorded voice reading the questions but record their answers using a digital device (O'Reilly, Hubbard, Lessler, Biemer, and Turner 1994). ACASI modes have been found to provide data quality comparable with traditional self-administered surveys with regard to reporting of sensitive behavior (Couper, Singer, and Tourangeau 2003). Furthermore, respondents seem to report more socially undesirable answers in ACASI compared with face-to-face surveys (Tourangeau and Smith 1996; Tourangeau and Yan 2007). A third way of humanizing web surveys is to incorporate both auditory and visual human characteristics in an online survey. This can be done by including videos of real interviewers, making it a dynamic nonresponsive feature (Fuchs and Funke 2007; Gerich 2008), or virtual agents, which are both dynamic and responsive (Lind, Schober, Conrad, and Reichert 2013; Conrad, Schober, Jans, Orlowski, Nielsen, et al. 2015). Little is known about response behavior with features that include both auditory and visual human characteristics in web surveys (Tourangeau et al. 2013). The few studies of response behaviors in humanized web surveys have mainly focused on the disclosure of sensitive information (e.g., Gerich 2008; Von der Pütten, Hoffmann, Klatt, and Krämer 2011) or on comprehension and engagement measured by response accuracy and interviewer-respondent verbal and nonverbal interactions (Conrad et al. 2015).

Intrigued by the possibility of improving data quality of web surveys by incorporating human-like features that contain both auditory and visual stimuli, we developed a "video-web" mode. In this mode, respondents see prerecorded video clips of an interviewer reading the questions and then record their own answers. Compared with the use of pictures and a human voice in an online survey, the human-like features in a video-web survey, which are dynamic but nonresponsive, are much more prominent as they contain moving images and the sound of an interviewer. Although some recent work has focused on virtual agents that are dynamic and responsive (e.g., Lind et al. 2013; Conrad et al. 2015), we believe the video-web mode deserves further exploration. First, most of the research on video-web modes has been conducted with convenience samples consisting only of students. Second, field studies investigating video-web are rare; most studies are conducted in lab settings, which makes

people aware of being in an experiment (Tourangeau et al. 2003). Lastly, the disclosure of sensitive information has been given the most attention in the past studies, while engagement, which may be stronger in video-web surveys than in traditional web surveys, has not.

For the field study reported here, we used a multistage cluster sampling method, using an address sample; address-based sampling is a method that generates representative samples (Link, Battaglia, Frankel, Osborn, and Mokdad 2008). Like the prior studies, we examine whether using a video-web mode elicits more social desirability bias by measuring the disclosure of sensitive information and compare these outcomes to a traditional web survey and to CAPI and CATI. Then, we investigate how engaged respondents are by studying indicators of satisficing behavior in the video-web mode and the traditional web survey. We focus on don't know (DK) answers and primacy and recency behavior. Finally, we compare primacy and recency behavior in video-web to CATI.

This paper proceeds as follows. In section 2, we start by providing what is known about response behavior in video-web surveys based on previous studies, followed by section 3, which presents the hypotheses of the current study. In section 4, we present our design. Then, our results follow in section 5. Finally, section 6 contains our conclusions and discusses possibilities for future work.

2. SUMMARY OF PREVIOUS STUDIES ON VIDEO-WEB SURVEYS

What we call "video-web" mode is dynamic, contains visual as well as aural stimuli, and is nonresponsive (i.e., there is no interaction between the respondent and the video-interviewer). This mode is created by recording an interviewer, who reads the questions. These videos are then incorporated into the web survey (see also section 4.3 for more design information).

Studies focusing on video-web surveys are limited thus far (see Table 1). Most work concerned with human-like features in web modes has concentrated on the disclosure of sensitive information relative to traditional web surveys and interviewer-administered modes (e.g., Fuchs and Funke 2007; Gerich 2008). These studies have mostly been conducted in laboratory settings with student samples. Krysan and Couper (2003) were the first to implement video recordings of interviewers in a web survey. Their laboratory-based findings show that interviewer race effects operate similarly for video-recorded and live interviewers with race-related questions. In addition, Fuchs (2009) found interviewer gender effects for video-web surveys similar to those reported in studies of face-to-face interviews (e.g., Tu and Liao 2007). Still, the results of Fuchs's study were somewhat mixed in its comparisons of video-web with traditional web; same-gender as well opposite-gender effects were found for questions related to relationships and sexual behaviors.

Table 1. Video-Web Survey Studies (Presented in Chronological Order)

Study	Participants	Size (*n*)	Contact mode	Response mode	Human-like feature	Main outcomes
Krysan and Couper (2003)	Lab study: white and African American subjects from Ann Arbor/Ypsilanti, MI, USA	160	F2F, newspaper advertisement, posters, church and community group visits	CAPI, video-web (random allocation)	Video of real interviewers, dynamic, nonresponsive	Interviewer's race operates similarly for video-recorded interviewers and live interviewers; hence interviewer effects are shown for video-web.
Fuchs and Funke (2007)	Field study: University of Kassel students, Germany	398	E-mail invitation	Traditional web, video-web (random allocation/ traditional web for second part survey)	Video of real interviewers, dynamic, nonresponsive	Unit nonresponse rate and break-off rate higher in video-web than traditional web. No difference in social desirability bias. Scores on social presence feelings higher for traditional web than video-web.
Gerich (2008)	Lab study: University of Linz students Austria	200	F2F in public area of the main building of the university	PAPI, F2F, ACASI, video-web (random allocation/PAPI for second part survey)	Video of real interviewers, dynamic, nonresponsive	Higher reported norm deviation in video-web, effect was mainly due to a low score in ACASI. There are some indications that video-web is less contaminated with socially desirable answering behavior.
Fuchs (2009)	Field study: Online access panel, University of Kassel students, Germany	880	E-mail invitation	Traditional web, video-web (random allocation)	Video of real interviewers, dynamic, nonresponsive	Interviewer gender effects seem to occur similarly for video-web compared with what is previously reported on face-to-face interviews.

Note—CAPI is computer-assisted personal interviewing; PAPI is paper-and-pencil interviewing; ACASI is audio computer-assisted self interviewing; F2F is face-to-face interviewing.

Fuchs and Funke (2007) were the first to compare a video-web survey to a traditional web survey. They found no differences between the two web modes with respect to disclosure of sensitive behavior. Furthermore, respondents indicated that they felt less personally addressed in the video-web survey than in the traditional web survey and that their traditional web answers were more similar to answers they would have (hypothetically) given in a face-to-face survey than the answers in the video-web survey. Gerich (2008) conducted a broader mode comparison study comparing a video-web mode with a paper and pencil questionnaire (PAPI), a face-to-face interview (F2F), and an ACASI interview. Although only small mode effects were observed, the video-web mode showed promising results regarding the disclosure of sensitive information compared with the other modes.

3. HYPOTHESES FOR THE CURRENT STUDY

How response behavior in video-web differs from traditional web and interviewer-administered surveys still remains unclear. Therefore, we tested hypotheses with respect to the disclosure of sensitive information and respondents' engagement in video-web compared with three other survey modes—traditional web, CAPI, and CATI.

Our first two hypotheses focus on the number of socially desirable responses by respondents. It is known that in interviewer-administered modes socially desirable responding is more likely to occur because interviewers influence respondents by their behavior, personal appearance, or voice characteristics, and the interaction may also affect respondents. A consistent finding in mode comparison studies is that in traditional web surveys respondents are inclined to answer (especially) sensitive questions more truthfully than in interviewer-administered modes (e.g., Greene, Speizer, and Wiitala 2008; Heerwegh and Loosveldt 2008; Heerwegh 2009). However, in comparing web and interviewer-administered modes, it is not clear whether effects are due to behavior in the interaction (which may be affected by proper interviewer training) or due to personal appearance and voice characteristics. Regardless of how an interviewer is trained, the interviewer's appearance may prime respondents: visual attributes such as skin shade, hair style, clothing, and so on and voice characteristics such as accent, pitch, prosody, and vocal quality will affect respondents' judgments about the interviewer's "social identity"—their gender, race, social class, education, or religious affiliation (Conrad et al. 2015).

Thus, although a video-web survey is nonresponsive, it may elicit socially desirable responses because of the interviewer's appearance. For instance, a video-web survey that contains a middle-class, middle-aged woman with no prominent accent may be expected to bring certain social norms to respondents' minds. Because these norms (e.g., "it is your civil obligation to vote") are brought to mind, respondents may adjust their answers to become more

consistent with the norm. So when respondents look at the recorded interviewer in video-web, their feeling of anonymity may be reduced. The interviewer's appearance may also prime certain attitudes, resulting in more socially desirable responses than in traditional web. However, compared with interviewer-administered modes, interviewer presence is subtler in video-web. Hence, it can be expected that socially desirable responding occurs more often in interviewer-administered modes. We present the following hypotheses about the disclosure of sensitive information:

Hypothesis 1a: The number of socially desirable responses will be higher in video-web than in a traditional web survey.

Hypothesis 1b: The number of socially desirable responses will be lower in video-web than in interviewer-administered modes (CAPI and CATI).

With respect to satisficing, our study examines the number of "don't know" (DK) answers, presumed to be a manifestation of strong satisficing (Vannette and Krosnick 2014). Whether a DK response is a "true" answer, an indicator of satisficing, or a socially desirable response is hard to determine. Items can be labeled "sensitive" or "difficult," but these labels interact with respondent characteristics as well (Stocké 2006). For example, a sensitive item on voting may only be sensitive for people who did not vote. While some may argue that a DK answer can be a socially desirable response to a sensitive item (Turner and Martin 1984), Shoemaker, Eichholz, and Skewes (2002) argue that DKs are more related to question difficulty than to question sensitivity. According to Shoemaker and colleagues (2002), respondents who provide DK responses start cognitively processing the question and then realize the question is too hard to answer, resulting in a DK response. In this study, we believe it is likely that DK responses, apart from reflecting an actual "true" answer, represent a form of satisficing; that is, people may give a DK response to avoid (further) thinking. We selected all questions that contained a DK option. This selection included attitudinal items that may be sensitive, but also questions on behaviors and personal characteristics for which it is less likely that respondents would provide a DK response because of question sensitivity (Shoemaker et al. 2002).

Although there are no studies published in which satisficing behavior is investigated directly in video-web surveys, the literature provides some indirect support for the hypothesis of less satisficing behavior in video-web than in traditional web surveys. Subtle human cues, as those implemented in video-web surveys, may improve the level of engagement because the video-recorded interviewer may hold respondents' attention, stimulating them to answer survey questions carefully (Reeves and Nass 1997; Tourangeau et al. 2003; Tourangeau et al. 2013). This may result in more substantive answers than in a traditional web survey. Because DK responses were not explicitly offered in CATI and CAPI, DK behavior in the video-web mode was not compared with these interviewer-administered modes. Based on these arguments, we formulated the following *hypothesis* on DK responding as an indicator of satisficing:

Hypothesis 2: The number of DK answers will be lower in video-web than in traditional web.

A second indicator of satisficing is the tendency for respondents to select the first response option they consider (a "primacy" effect) or to select one of the last options (a "recency" effect; Vannette and Krosnick 2014). For items with ordinal response options, primacy effects are likely to occur when the response options are presented visually and recency effects are likely when the response options are presented aurally. In both video-web and traditional web, the response options are presented visually, but in video-web the response options are also presented aurally. Therefore, we still expect a difference between these web modes relating to primacy and recency. Compared with CATI, in which response options are presented aurally, it can be expected that primacy effects are more likely to occur in video-web while recency effects are more likely to occur in CATI. Since in CAPI show cards were used, this could have affected primacy and recency behavior. Therefore, CAPI was excluded from this analysis. Hence, the following hypotheses were formulated related to primacy and recency as indicators of satisficing:

Hypothesis 3a: Primacy effects will be smaller in video-web than in traditional web, while recency effects will be smaller in traditional web than in video-web.

Hypothesis 3b: Primacy effects will be larger in video-web than in CATI, while recency effects will be larger in CATI than in video-web.

4. METHODS

4.1 Data Collection

To compare response behavior across survey modes, we did an experiment in an additional round of the Dutch European Social Survey (ESS). We used a slightly modified version of the ESS round 5 core questionnaire sections on politics (41 items), subjective well-being (36 items), and social demographics (53 items). Data were collected from March to June 2012 in the Netherlands by GfK Panel Services Benelux. Multistage cluster sampling was used. In the first stage, 40 municipalities (out of a total of 441 municipalities) were selected with equal probability from within the 12 Dutch provinces and from areas at different urbanization levels (i.e., highly urbanized to rural areas). Then, in the second stage, address-based sampling was applied; addresses were randomly selected from databases containing all the addresses from the ZIP codes within these 40 municipalities. The total sample consisted of 3,496 households. All selected households received a letter in which the goal of the survey was introduced and their selection for this study was explained. A form of the last-birthday method was used to select one individual within each selected household (when more than one individual lived in the household). This sampling

Table 2. Response Rates and Cooperation Rates per Experimental Group

	Group 1	Group 2	Group 3	Group 4	Group 5
Contact Mode	Face-to-face Choice of CAPI or web	Telephone Choice of CATI or web	Telephone CAPI	Telephone CAPI	Telephone Web
RR1	54.9	34.8	20.8	31.3	36.7
COOP1	60.6	40.6	23.6	32.8	41.4

method, which is standard in the ESS, requires that the interviewer asks which person in the household had his or her birthday closest to a randomly chosen date. The identified individual was selected for the survey. Of the 3,496 households selected for this study, 824 households participated. The overall response rate was 37.5 percent, and the cooperation rate 42.1 percent (American Association for Public Opinion Research 2011, RR1 and COOP1).

4.2 Experimental Design

The participating households were randomly allocated to one of five experimental groups. An embedded experiment investigated the effect of mode choice on participation. Group 1 was contacted face-to-face and could choose between a personal interview (CAPI) and a web survey. Group 2 was contacted by telephone and could choose between a telephone interview (CATI) and a web survey. Households in the other three groups were contacted by telephone and then randomly allocated to CAPI (group 3), CATI (group 4), or web (group 5). Table 2 shows the response rates and cooperation rates for these groups. The number of households per group is displayed in Appendix A of the online Supplementary Materials.

With this design, selection bias may occur for two reasons. First, sampled households were contacted either face-to-face or by telephone, and this could have affected the households that were reached by the different contact modes. Second, choosing a response mode may have affected the composition of respondents in a group as different respondents may choose different response modes. Because the contact mode and response modes offered in group 1 (face-to-face and a CAPI/web choice) differed from the contact mode in groups 2–5 (telephone) and differed from the mode choice offered in group 2 (CATI/web), we decided to exclude group 1 from the study. The four remaining groups, containing 536 respondents, were compared using demographic variables to reveal possible selection bias due to mode choice (i.e., self-selection) and random allocation (see Table 3 in Results section). Note that due to the internet penetration rate of 95 percent of all households, in the Netherlands web surveys do not necessarily yield a coverage bias (Deutskens, Ruyter, Wetzels, and Oosterveld 2004).

Table 3. Composition of Respondents in Experimental Groups 2–5 ($n = 536$)

	CATI choice ($n = 100$)	Web choice ($n = 125$)	CAPI random ($n = 100$)	CATI random ($n = 106$)	Web random ($n = 105$)
Gender					
Male	39.0	40.8	48.0	49.1	53.3
Female	61.0	59.2	52.0	50.9	46.7
Age					
Young/middle (18–55)	36.0	62.4	48.0	47.2	60.0
Senior (>55)	64.0	37.6	52.0	52.8	40.0
Education					
Low	67.3	63.6	70.7	75.2	61.9
High	32.7	36.4	29.3	24.8	38.1

4.3 Design Video-Web Mode

The video-web mode consisted of prerecorded clips of an interviewer reading the questions to the respondent (see figure 1). Since most interviewers in the CAPI and CATI mode were white females, for the video-web condition, a white female interviewer was recorded for the entire interview. Each item was presented on a new screen, and the video automatically started when respondents entered the screen. The question was not repeated in writing on the screen, but the respondent could watch every clip as often as desired. As the question was read by the video-interviewer, response options were presented on the screen. It was possible to select response options or change selections while the video-interviewer was reading the question and once the clip stopped playing. To make the video-web mode more comparable with the interviewer-administered modes (CAPI and CATI), for some questions the video-interviewer read the response options out loud as well (see section 4.4 on primacy and recency). When more explanation for a question was necessary (e.g., "Only one answer possible, select the most important."), this was displayed on the screen.

Although the respondent was able to see the interviewer in the video, no interaction was possible between the interviewer and respondent. Therefore, the video-web mode presents respondents with a dynamic, nonresponsive human-like feature. Giving an answer was required to continue with the next survey question. A DK response option was included between brackets for every item (as in the traditional web survey). Half of the respondents who selected or were randomly allocated to the web questionnaire started with a traditional web survey and then switched to the video-web mode, the other half started with video-web and then switched to the traditional web mode. The web respondents were not aware of these two web modes prior to the switch, and which mode they received first was determined by a random procedure. The web mode switch took

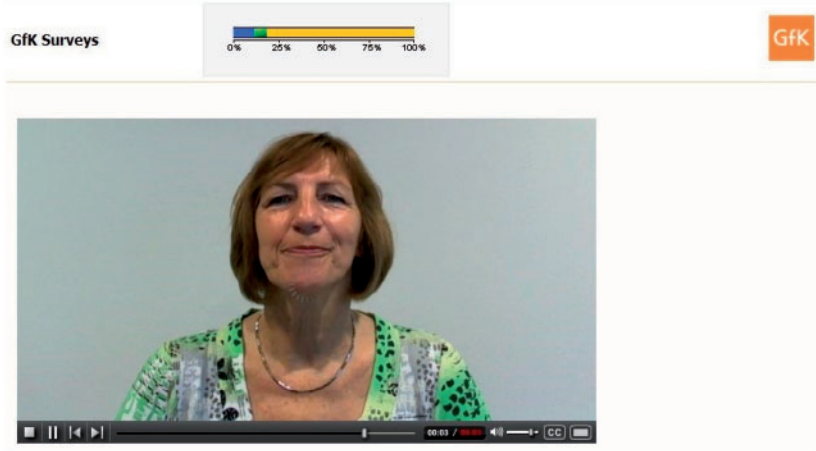

Figure 1. Video-Web Display.

place after part B of the ESS questionnaire. By letting the web respondents experience both web modes, we were able to conduct within-group as well as between-group analyses regarding response behavior in these two web surveys. Which questions of the ESS questionnaire were asked before and after the web mode switch can be found in Appendix A of the online Supplementary Materials.

4.4 Dependent Variables

Our primary goal was to compare a video-web survey to a traditional web survey in terms of satisficing behavior and the disclosure of sensitive information. In our study, satisficing includes DK responding and primacy and recency behavior. Regarding the disclosure of sensitive information, we focused on possible social desirability bias among the web modes. We also analyzed differences of possible socially desirable responding and primacy and recency between the video-web mode and interviewer-administered modes. DK answers in the video-web mode were not compared with the interviewer-administered modes because a DK option was not explicitly given in CAPI and CATI.

4.4.1 DK responding.

Within our version of the ESS questionnaire, 117 items offered a DK option (see items in Appendix A.1 of the online Supplementary Materials). The questions on politics ($n = 41$) and subjective well-being ($n = 36$) all contained a DK, and 40 out of the 53 questions on social demographics had a DK option. Thus, 41 items with a DK option were presented before the web mode switch and 76 items with a DK option afterwards. For each respondent, we removed the items that were not applicable and calculated the mean percentage of DK responses given by the respondent.

4.4.2 Primacy and recency.
Primacy and recency effects were measured for items with ordinal response options (see items in Appendix A.2 of the online Supplementary Materials.). In total, 20 items were analyzed, nine items on politics presented before the mode switch and 11 items on subjective well-being presented after the mode switch. For all of these items, the response options were read out loud in the video-web mode and the interviewer-administered modes. To measure primacy, we first created a binary indicator that took on a value of 1 if one of the first response options was selected and a value of 0 if any other answer was provided. For three-point scales, only the first response option was coded as 1; for four-point and five-point scales, the first two response options were coded as 1. To measure recency, we first created a binary indicator that was coded as 1 if one of the last response options was selected and 0 if any other answer was provided. For three-point scales, only the last response option received a value of 1; for four-point and five-point scales, the last two response options received a value of 1. Then, we calculated the mean percentage of primacy and recency behavior for each respondent.

4.4.3 Socially desirable responding.
To find out whether socially desirable responding differed across modes, items were selected that have been shown to be sensitive to socially desirable answering behavior or have widely shared social desirability connotations (see items in Appendix A.3 of the online Supplementary Materials.), such as income and religious service (Jäckle, Roberts, and Lynn 2006, 2010). We included other topics about which respondents might be hesitant to disclose their true opinions, such as immigration and equality between the sexes. In total, we selected 19 items, nine items on politics presented before the mode switch and eight items on subjective well-being and two items on income presented after the mode switch. When necessary, items were rescaled so that the response option with the highest number on the scale was the most socially desirable answer. For some of the items, the middle response options were more likely to be the socially desirable answer (religiosity, church attendance, and income). To measure possible socially desirable responding, the number of sensitive admissions was counted and this score was used for mode comparisons.

5. RESULTS

5.1 Composition of the Experimental Groups

Due to the design of the study, the apparent effects of mode on answering behavior cannot be interpreted as causal effects because they may also reflect differences between the respondents of the different modes. To analyze the

impact of self-selection and randomization, we studied the gender, age, and education levels of the respondents in experimental groups 2–5 (see Table 3).

The respondents in experimental group 2 ($n = 225$), who could choose between CATI or web administration, are presented separately (i.e., CATI choice and Web choice) because these two groups were also separated in the mode comparison analyses. We focused on the respondents' gender, age, and level of education because these are likely indicators of response mode preferences (Loges and Jung 2001; Schneider, Cantor, Malakhoff, Arieira, Segel, et al. 2005). This set of variables is still somewhat limited. As a result, it is difficult to draw causal conclusions from our analysis.

For gender, we found no significant differences between all five modes (χ^2 (4, $n = 536$) = 6.16, $p = 0.18$, $\omega = 0.11$). However, more female respondents participated in the experimental groups with mode choice (χ^2 (1, $n = 536$) = 5.47, $p = 0.02$, $\omega = 0.10$). For age, significant differences were found; more senior respondents participated in the CATI modes, and more younger and middle-aged respondents participated in the web modes (χ^2 (4, $n = 536$) = 19.88, $p = 0.001$, $\omega = 0.19$). For level of education, no significant differences were found (χ^2 (4, $n = 520$) = 5.57, $p = 0.23$). Based on these results, gender and age were included as covariates in the analysis.

5.2 Effects of Web Modes on Answering Behavior

We first present our findings on the differences in answering behavior between video-web and traditional web. In these analyses, web modes with choice are merged with the no-choice web modes since we found no significant differences between these groups. We included only those responses that were given before the mode switch to make a clean between-subjects comparison between the two web modes. We included gender and age in the analysis as covariates.

5.2.1 Socially desirable responding (H1a).

Analysis of covariance (ANCOVA) was used to examine the effect of mode on socially desirable responding by respondents. Mode was the independent variable, and the number of socially desirable responses as the dependent variable. Contrary to hypothesis 1a, after accounting for the effects of gender and age, there was no significant effect of mode on the number of sensitive admissions: F (1, 226) = 0.02, $p = 0.90$ (video-web M = 4.8, SE = 0.22 versus traditional web M = 4.8, SE = 0.22).

5.2.2 Don't know answering behavior (H2).

ANCOVA was also used to compare the number of DK answers in the video-web mode with the traditional web mode. Mode was included as the independent variable, and the mean percentage of DK responses as the dependent variable. ANCOVA indicated that the effect of mode on DK responding was not statistically significant: F (1, 226) = 0.547, $p = 0.46$ (video-web M = 3.7,

SE = 0.66 versus traditional web M = 3.0, SE = 0.65). Therefore, hypothesis 2 was not supported.

5.2.3 Primacy and Recency (H3a).
Another ANCOVA compared the number of primacy and recency responses in video-web and traditional web. Mode was included as the independent variable, and the mean percentage of primacy or recency responses was the dependent variable. For primacy, the analysis yielded no significant effect of mode on response behavior: F (1, 226) = 0.333, p = 0.98 (video-web M = 58.6, SE = 2.10 versus traditional web M = 58.5, SE = 2.06). For recency, there was also no significant effect of mode on response behavior: F (1, 226) = 0.536, p = 0.47 (video-web M = 22.0, SE = 1.74 versus traditional web M = 23.8, SE = 1.71). Hypothesis 3a was not supported.

5.3 Effects of Mode Switch on Answering Behavior

This section describes how the mode switch between video-web and traditional web affected respondents' answering behavior. In the repeated measures analyses of variance (ANOVA), the two sets of items—the set before the web mode switch and the set after the switch—were included as within-subject factors.

5.3.1 Socially desirable responding (H1a).
A repeated measures ANOVA was used to compare the effect of switching between the two web modes on socially desirable responding. The ANOVA results showed that there was no significant difference between the items that were presented before the mode switch and the items that were presented after: F (1, 228) = 1.878, p = 0.17. The video-web and traditional web respondents did not significantly change their answering behavior (video-web before the switch M = 4.8, SE = 0.22 versus video-web after M = 4.9, SE = 0.17; traditional web before M = 4.8, SE = 0.22 versus traditional web after M = 5.1, SE = 0.16).

5.3.2 DK responding (H2).
A repeated measures ANOVA used to compare the effect of switching between the two web modes on don't know responding yielded no significant effect of mode switch (F (1, 228) = 0.612, p = 0.44). The video-web and traditional web respondents did not significantly change their answering behavior (video-web before M = 3.7, SE = 0.66 versus after M = 2.5, SE = 0.62; traditional web before M = 3.0, SE = 0.65 versus after M = 3.4, SE = 0.61).

5.3.3 Primacy and recency (H3a).
We also used a repeated measures ANOVA to compare the effect of switching between the two web modes on primacy and recency. For primacy, the

ANOVA indicated that there was a significant effect (F (1, 228) = 140.3, $p < 0.001$; the mean of the set of items administered before the mode switch did significantly differ from the mean of the items after the switch (before switch M = 58.6, SE = 1.54; after switch M = 37.96, SE = 0.82), but there was no interaction between switch and mode (F (1, 228) = 0.062, $p = 0.80$; video-web before M = 58.6, SE = 2.01, versus after M = 38.5, SE = 1.17; traditional web before M = 58.5, SE = 2.06, versus after M = 37.4, SE = 1.15). For recency, a significant effect was found between the set of items before the mode switch and the set of items after the mode switch: F (1, 228) = 70.0, $p < 0.001$ (before switch M = 22.93, SE = 1.24; after switch M = 35.6, SE = 0.80). But there was also no interaction between switch and mode: F (1, 228) = 0.387, $p = 0.54$ (video-web before M = 22.0, SE = 1.74, versus after M = 33.9, SE = 1.15; traditional web before M = 23.8, SE = 1.71, versus after M = 37.3, SE = 1.13).

5.4 Additional Mode Comparisons

In this section, we compare the video-web mode to the interviewer-administered modes. The choice versus no-choice modes were again merged (because analyses comparing the choice and no-choice groups yielded no significant differences), and the analysis of socially desirable responding includes three survey modes: video-web, CATI, and CAPI. Since show cards were used in CAPI, this could have affected primacy behavior. Therefore, for primacy, video-web was compared to CATI only (CAPI was excluded from this set of analysis). We included only those responses that were given before the mode switch to make a clean between-subject comparison. Possible effects of gender and age are accounted for in the analyses.

5.4.1 Socially desirable responding (H1b).

ANCOVA was used to compare socially desirable responding in three different modes: video-web, CAPI, and CATI. The findings showed that there was a significant difference across the modes: F (2, 414) = 6.01, $p = 0.003$ (video-web M = 4.8, SE = 0.21; CATI M = 5.3, SE = 0.16; CAPI M = 5.8, SE = 0.22). Higher means indicate higher levels of socially desirable responding. Hypothesis 1b was supported.

Separate analyses were conducted for the nine sensitive items on which our overall index was based. The analyses per item showed significant or marginally significant differences across modes for seven of the items: one item on life satisfaction (B24) and six items on immigration (B35, B36, B37, B38, B39, and B40). Table 4 displays the differences across the three mode groups for these seven items.

Table 4. Mode Comparison of Socially Desirable Responding Among Sensitive Items

	Video-web[a] M (SE)	CATI[b] M (SE)	CAPI[c] M (SE)
B24 life satisfaction	7.2 (0.14)[bc]	7.7 (0.10)[a]	7.7 (0.15)[a]
B35 allowing immigrants in NL of same race/ethnicity	2.7 (0.07)[c]	2.6 (0.05)[c]	2.9 (0.07)[ab]
B36 allowing few/many immigrants in NL	2.6 (0.07)[c]	2.6 (0.05)[c]	2.8 (0.07)[ab]
B37 allowing immigrants in NL from poor non-EU countries	2.4 (0.08)[c]	2.5 (0.05)	2.7 (0.08)[a]
B38 effect of immigrants on economy	4.9 (0.18)[bc]	5.9 (0.13)[a]	5.8 (0.19)[a]
B39 effect of immigrants on cultural life	5.5 (0.19)[bc]	6.5 (0.14)[a]	6.4 (0.19)[a]
B40 effect of immigrants on living in NL	5.4 (0.19)[b]	5.9 (0.14)[a]	5.7 (0.19)

NOTE.—Letters in superscript indicate which means are significantly different from each other.

5.4.2 Primacy and Recency (H3b).
We compared the number of primacy responses in two mode groups—video-web and CATI—using ANCOVA. ANCOVA indicated that the effect of mode on primacy was statistically nonsignificant: F (1, 315) = 0.021, $p = 0.89$ (video-web M = 59.5, SE = 2.03; CATI M = 59.8, SE = 1.50). However, for recency behavior, a significant difference was found between video-web and CATI: F (1, 315) = 4.608, $p = 0.03$ (video-web M = 22.0, SE = 1.71; CATI M = 26.6, SE = 1.26). Thus, hypothesis 3b was not supported for primacy, but it was supported for recency; the last response options were chosen significantly more often by CATI respondents than video-web respondents.

6. DISCUSSION

As web surveys are used more and more, questions are raised about how data quality of this mode can be improved. Since web respondents enter their responses in a system that is minimally interactive, their impression of the interview situation is mostly a solitary one, where responses are not monitored during the interview. Since respondents usually are not intrinsically motivated to participate in a survey due to the noninteractive (or nonresponsive, Tourangeau et al. 2013) character of web surveys, they may be less likely to be fully engaged and, consequently, display satisficing behavior. However, an advantage of the nonresponsiveness of web surveys is that respondents are less likely to adjust their answers to social norms (i.e., web surveys evoke less social desirability bias than interviewer-administered surveys). This contrast

between satisficing and social desirability bias led us to test the effects of a video-web survey. In video-web, a video is played for each question, with an interviewer reading the question. This adds a dynamic human-like feature to the survey, which is assumed to increase the level of respondents' engagement without substantially increasing their impressions of being judged on their answers, due to the nonresponsiveness of the human-like feature.

We tested several hypotheses comparing video-web with traditional web. Since half of the respondents started with a video-web survey and then switched to the traditional web survey and the other half started with traditional web and then switched to video-web, we conducted two analyses for every hypothesis in which video-web was compared with traditional web. The first analysis focused on a comparison between video-web and traditional web, the second on the effects of the web mode switch on response behavior. We expected more socially desirable responding in the video-web mode compared with the traditional web mode. This hypothesis was not supported; no significant differences were found between video-web and traditional web. We also examined whether, compared with traditional web, there would be less satisficing behavior (primacy and recency effects and DK answers) in video-web. We found no differences between video-web and traditional web for don't know responding and primacy and recency. We carried out a further test of these hypotheses based on within-subject analyses. These analyses revealed that switching modes did not affect socially desirable responding or satisficing behavior of respondents in the two web modes. Overall, it seems that despite the visual and auditory representation of a human interviewer in a video-web mode, video-web is experienced by respondents much like a traditional web mode. This may be mostly due to the fact that video introduced a dynamic humanizing cue, but not a responsive one. Thus, in video-web, respondents are still fully aware of the fact that they are responding in private; there is no person on the other side who is monitoring their engagement, and there is no social sanctioning on satisficing behavior. Future studies should include debriefing questions concerning respondent experiences with the mode in which they answered the items and completion times per item to gather more information about web survey participation.

Additional hypotheses were formulated to test differences between video-web and interviewer-administered modes. First, we tested hypothesis 1b; we expected less socially desirable responding in video-web compared with CAPI and CATI. This hypothesis was supported. Studying the items separately, we found (marginal) significant differences for seven out of the nine sensitive items that were used for the analysis. Significant differences were mainly found between video-web and the interviewer-administered modes, though for some questions we found a larger difference between video-web and CAPI, and for other questions between video-web and CATI. In addition, we found significant differences between CAPI and CATI for two items. Analyses of these items showed that respondents gave more socially desirable answers in

CAPI than in CATI. These results indicate that video-web surveys do not provoke the same response behavior as in CAPI or CATI, while previous research suggests that response behavior in video-web may be similar to face-to-face surveys (Krysan and Couper 2003; Fuchs 2009). In the current study, we did not find such evidence. However, our questions differed from the questions used in those previous studies. It is possible that questions that are more gender-sensitive (e.g., items on sexual behaviors, Fuchs 2009) or race-sensitive (e.g., items on differences between black and white people, Krysan and Couper 2003) are more likely to provoke response behaviors in video-web that are similar to face-to-face interviews (i.e., gender effects and race effects). Apart from the fact that other questions were used in the current study, we also did not record more than one interviewer. Therefore, we could not investigate differences in response behavior related to gender or race.

We also examined primacy and recency effects. We found no differences in primacy behavior between video-web and CATI, but we did find significantly more recency behavior in CATI. That CATI is an aural mode may explain this difference in recency effects. In web, respondents can decide for themselves how long they spend on a particular question, while in CATI interviewers decide the pace of the interview when they aurally present the questions and response options.

Our results extend the body of literature on video-web surveys and more generally on human-like features in online surveys. This study provides insights in how video-web is received in a field study with a more diverse population than the earlier studies. Like Fuchs and Funke (2007), we found no difference in the disclosure of sensitive information between the web modes. Furthermore, no differences in DK responding and primacy and recency behavior were found. We argue that for a humanizing feature to fully increase the level of engagement it would require responsiveness. Responsive humanizing cues may give respondents the impression they are interacting with a system that may also judge their answers. Such an impression is possible with virtual agents (Conrad et al. 2015) or with video recordings of a real interviewer with possible interviewer reactions. Creating responsive video-web surveys may take some time because interviewer actions for all possible respondents' actions need to be recorded, and software needs to be developed that should be able to show the video clip that fits best in a certain situation (Tourangeau et al. 2013). However, specific interviewer actions could be recorded in order to, for example, decrease satisficing behavior. For example, after a respondent has selected "don't know," a video may be played with an interviewer saying "I have noticed that you selected 'don't know,' but could you please consider once again if this is your final answer?" Note that in such a video recording the phrase "I have noticed" refers to the fact that in video-web respondents are entering responses in the interface via keyboard or mouse clicks. Such a phrase would be odd to use in an interview with a real interviewer where respondents provide their answers aurally, and, as such, "responsive" video recordings are

Table 5. Virtual Agents in Survey Studies (Presented in Chronological Order)

Study	Participants	Size (*n*)	Contact mode	Response mode	Human-like feature	Main outcomes
von der Pütten, Hoffmann, Klatt, and Krämer (2011)	Lab study: University of Duisburg-Essen students, Germany	81	General advertising on campus	Video-web (PAPI for demographic items)	Virtual agent, dynamic, responsive	Talkative virtual agent affected responses; respondents provided more personal information.
Lind, Schober, Conrad, and Reichert (2013)	Lab study: subjects from New York City area, NY, USA	235	Online advertising on Craig's List, print and online ads in *Village Voice*	CAPI, ACASI, video-web (random allocation/ PAPI for debriefing questions)	Virtual agent, dynamic, responsive	Virtual agents with limited human features can reduce the disclosure of sensitive information. Sensitive responses to virtual agents fall between ACASI and CAPI.
Conrad, Schober, Jans, Orlowski, Nielsen, et al. (2015)	Lab study: subjects from Ann Arbor area, MI, USA	73	Online advertising on Craig's List and word of mouth	Video-web (random allocation/traditional web for debriefing questions)	Virtual agent, dynamic, responsive	Dialog capability and facial movement of virtual agents differently affect respondents' comprehension and feelings of engagement with the virtual agent.

still quite different from the real conversation in interviewer-administered surveys. Still, there are many possibilities for responsive video recordings to mirror interview conversation. For example, a responsive video-web survey with interviewers explaining definitions would also add a more conversational interviewing aspect (Conrad and Schober 1999).

Much remains to be learned about humanizing web surveys and the conditions in which human-like cues actually improve the data quality of web surveys. Recent studies have developed different versions of virtual agents (see Table 5). Von der Pütten, Hoffmann, Klatt, and Krämer (2011) manipulated the wordings of a virtual agent by creating a talkative and nontalkative interviewer and varied the extent to which the virtual agent disclosed personal information to the respondent during the introduction of the interview. They found that talkative interviewers received more personal information from respondents and respondents themselves became more talkative as well when replying to open-ended items. Lind and colleagues (2013) studied facial expressions of a virtual agent; they developed a high-animation virtual agent (i.e., more human-like) and a low-animation virtual agent (i.e., more robot-like). No differences in response behavior were found between the two virtual agent conditions. Conrad and his colleagues (2015) tested how dialog capability and facial movement of virtual agents affected respondents' question comprehension and engagement with the survey interview. Their findings indicate that high dialog capability of the virtual agent positively affects respondents' comprehension; more accurate responses were provided to the questions. Virtual agents with the most facial movement yielded a higher level of respondents' engagement with the interview. However, the virtual agents with the most facial movement were also rated "less natural" than the low facial virtual agents. For future studies, we suggest further exploring the possibilities of human-like features in web surveys, taking into account the responsiveness and humanness of the virtual interviewer and the resulting effects on survey participation and cooperation.

Supplementary Materials

Supplementary materials are available at Journal of Survey Statistics and Methodology online.

References

American Association for Public Opinion Research (2011), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates For Surveys*, Oakbrook Terrace, IL: AAPOR.

Callegaro, M., R. P. Baker, J. Bethlehem, A. S. Goritz, J. A. Krosnick, and P. J. Lavrakas (2014), *Online Panel Research: A Data Quality Perspective*, Hoboken, NJ: Wiley.

Conrad, F. G., and M. F. Schober (1999), "Conversational Interviewing and Data Quality [online]," Available at: http://www.bls.gov/ore/pdf/st990250.pdf.

Conrad, F. G., M. F. Schober, M. Jans, R. A. Orlowski, D. Nielsen, and R. Levenstein (2015), "Comprehension and Engagement in Survey Interviews With Virtual Agents," *Frontiers in Psychology*, 6, 1578.

Couper, M. P., E. Singer, and R. Tourangeau (2003), "Understanding the Effects of Audio-CASI on Self-Reports of Sensitive Behavior," *Public Opinion Quarterly*, 67, 385–395.

Deutskens, E., K. Ruyter, M. Wetzels, and P. Oosterveld (2004), "Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study," *Marketing Letters*, 15, 21–36.

Fricker, S., M. Galesic, R. Tourangeau, and T. Yan (2005), "An Experimental Comparison of Web and Telephone Surveys," *Public Opinion Quarterly*, 69, 370–392.

Fuchs, M. (2009), "Gender-of-Interviewer Effects in a Video-Enhanced Web Survey: Results From a Randomized Field Experiment," *Social Psychology*, 40, 37–42.

Fuchs, M., and F. Funke (2007), "Video Web Survey—Results of an Experimental Comparison with a Text-Based Web Survey," in *The Challenges of a Changing World*, ed. M. Trotman, Proceedings of the fifth International Conference of the Association for Survey Computing, Berkeley, CA, Association for Survey Computing, pp. 63–80.

Gerich, J. (2008), "Real or Virtual? Response Behavior in Video-Enhanced Self-Administered Computer Interviews," *Field Methods*, 20, 356–376.

Greene, J., H. Speizer, and W. Wiitala (2008), "Telephone and Web: Mixed-Mode Challenge," *Health Services Research*, 43, 230–248.

Heerwegh, D. (2009), "Mode Differences Between Face-to-Face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects," *International Journal of Public Opinion Research*, 21, 111–121.

Heerwegh, D., and G. Loosveldt (2008). "Face-to-Face Versus Web Surveying in a High-Internet Coverage Population: Differences in Response Quality," *Public Opinion Quarterly*, 72, 836–846.

Holbrook, A. L., M. C. Green, and J. A. Krosnick (2003), "Telephone vs. Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias," *Public Opinion Quarterly*, 67, 79–125.

Jäckle, A., C. Roberts, and P. Lynn (2006), "Telephone Versus Face-to-Face Interviewing: Mode Effects on Data Quality and Likely Causes. Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project." ISER Working paper 2006-41, University of Essex, Colchester, UK.

——— (2010). "Assessing the Effect of Data Collection Mode on Measurement," *International Statistical Review*, 78, 3–20.

Krosnick, J. A. (1991), "Response Strategies for Coping With the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, 5, 213–236.

Krysan, M., and M. P. Couper (2003), "Race in the Live and the Virtual Interview: Racial Deference, Social Desirability, and Activation Effects in Attitude Surveys," *Social Psychology Quarterly*, 66, 364–383.

——— (2006), "Race of Interviewer Effects: What Happens on the Web?" *International Journal of Internet Science*, 1, 17–28.

Lind, L. H., M. F. Schober, F. G. Conrad, and H. Reichert (2013), "Why do Survey Respondents Disclose More When Computers ask the Questions?" *Public Opinion Quarterly*, 77, 888–935.

Link, M. W., M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad (2008), "A Comparison of Address-Based Sampling (ABS) Versus Random-Digit Dialing (RDD) for General Population Surveys," *Public Opinion Quarterly*, 72, 6–27.

Loges, W. E., and J. Jung (2001), "Exploring the Digital Divide: Internet Connectedness and Age," *Communication Research*, 28, 536–562.

Lozar Manfreda, K., M. Bosnjak, J. Berzelak, I. Haas, and V. Vehovar (2008), "Web Surveys Versus Other Survey Modes: A Meta-Analysis Comparing Response Rates," *International Journal of Market Research*, 28, 536–562.

O'Reilly, J. M., M. L. Hubbard, J. T. Lessler, P. P. Biemer, and C. F. Turner (1994), "Audio and Video Computer Assisted Self-Interviewing: Preliminary Tests of New Technologies for Data Collection," *Journal of Official Statistics*, 10, 197–214.

Reeves, B., and C. Nass (1997), *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*, Cambridge: CSLI and Cambridge University Press.

Schneider, S. J., D. Cantor, L. Malakhoff, C. Arieira, P. Segel, K. Nguyen, and J. G. Tancreto (2005), "Telephone, Internet and Paper Data Collection Modes for the Census 2000 Short Form," *Journal of Official Statistics*, 21, 89–101.

Shoemaker, P. J., M. Eichholz, and E. A. Skewes (2002), "Item Nonresponse: Distinguishing Between Don't Know and Refuse," *International Journal of Public Opinion Research*, 14, 193–201.

Sproull, L., M. Subramani, S. Kiesler, J. H. Walker, and K. Water (1996), "When the Interface is a Face," *Human-Computer Interaction*, 11, 97–124.

Stocké, V. (2006), "Attitudes Toward Surveys, Attitude Accessibility and the Effect on Respondents' Susceptibility to Nonresponse," *Quality & Quantity*, 40, 259–288.

Tourangeau, R., F. G. Conrad, and M. Couper (2013), *The Science of Web Surveys*, New York: Oxford University Press.

Tourangeau, R., M. P. Couper, and D. M. Steiger (2003), "Humanizing Self-Administered Surveys: Experiments on Social Presence in Web and IVR Surveys," *Computers in Human Behavior*, 19, 1–24.

Tourangeau, R., and T. W. Smith (1996), "Asking Sensitive Questions. The Impact of Data Collection Mode, Question Format, and Question Context," *Public Opinion Quarterly*, 60, 275–304.

Tourangeau, R., and T. Yan (2007), "Sensitive Questions in Surveys," *Psychological Bulletin*, 133, 859–883.

Tu, S. H., and P. S. Liao (2007), "Social Distance, Respondent Cooperation, and Item Nonresponse in Sex Survey," *Quality and Quantity*, 41, 177–199.

Turner, C. F., and E. Martin (1984), *Surveying Subjective Phenomena: Volume 2*, New York: Russel Sage Foundation.

Vannette, D. L., and J. A. Krosnick (2014), "Answering Questions: A Comparison of Survey Satisficing and Mindlessness," in *The Wiley-Blackwell Handbook of Mindfulness,* eds. A. Le, C. T. Ngnoumen, and E. J. Langer, pp. 312–327, Hoboken, NJ: Wiley-Blackwell.

Von der Pütten, A. M., L. Hoffmann, J. Klatt, and N. C. Kramer (2011), "Quid pro quo? Reciprocal Self-Disclosure and Communicative Accommodation Towards a Virtual Interviewer," Proceedings of the 11th International Conference on Intelligent Virtual Agents, Lecture Notes in Artificial Intelligence 6985, Berlin Heidelberg, Springer, pp. 183–194.