How Does Video Interviewing Affect the Interviewers’ and Respondents’ Paralinguistic Behaviors? A First Exploration

Hanyu Sun[1] & Frederick G. Conrad[2]

[1] *Westat*
[2] *University of Michigan*

Abstract
Video interviews have been gradually adopted by survey organizations as an alternative to in-person interviews as a mode of data collection. Recent studies have shown that video and in-person interviews elicited similar levels of respondents’ rapport with interviewers and similar quality data. However, little is known about whether the presence and prevalence of interviewer and respondent paralinguistic behaviors (e.g., disfluencies such as “*uh*” and “*um*” or laughter) vary between the two modes, and when they do, how might this affect survey outcomes? To address these questions, we coded the presence of six paralinguistic behaviors in 710 question-answer (Q-A) sequences in 15 in-person interviews and 12 video interviews conducted by professional survey interviewers in a laboratory experiment. Most of the paralinguistic behaviors occurred equally often in the two interviewing modes except laughter which was significantly more prevalent in video than in-person interviews. We attributed the increased laughter in video interviews as a nervous response to greater communication difficulties in that mode. Nonetheless, this did not differentially impact the prevalence of respondents’ adequate responses, indicating (indirectly) that data quality was equivalent in the two modes. These findings bolster the emerging narrative that when interviewed via video, respondents’ experience and their answers are very similar to when they are interviewed in person.

Keywords: live video interviews, in-person interviews, paralinguistic behaviors

Video interviewing has been increasingly adopted by survey organizations as the underlying technology has become more mature and widely available, and because the pandemic drove organizations to find alternative ways to safely collect data requiring face-to-face interaction. In both in-person and video interviews, interviewers and respondents can see and hear each other in real time, enabling interviewers to rely on visual information about respondents, such as their facial expressions, and to provide visual materials to respondents, such as showcards or standardized test materials. It is plausible that reading respondents' facial expressions and presenting visual material to them is generally no more difficult and sometimes easier via video than in person. For instance, a recent study showed that interviewers exhibited similar question reading behaviors in both in-person and video interviews and respondents provided a higher percentage of responses that matched one of the provided response categories in video than in-person interviews (Kokoska et al., 2024). Moreover, respondents in video interviews and their counterparts in in-person interviews reported similar levels of rapport with interviewers, disclosed similar amounts of sensitive information (Sun et al., 2020), exhibited similar levels of item nonresponse and straightlining, provided answers of similar length to open-ended questions, and had similar interview experiences (Endres et al., 2022). However, little is known about how communication between interviewers and respondents might be affected by video mediation and how this might impact the respondent's answers to a variety of questions.

Paralinguistic Behaviors in Spontaneous Speech

An indication of how well two conversational partners, including a survey interviewer and a respondent, are communicating can sometimes be found in their paralinguistic utterances. These are generally (but not always) non-lexical parts of speech that can serve as cues about the speaker's current state of mind including their understanding of the speaker and their affective state or mood (e.g., Brown, 1977; Hancock, 2004). Speakers produce *disfluencies*—"um" and "uh", hesitations or pauses, and repairs—to indicate planning activities or difficulty in word finding (Clarck & Fox Tree, 2002; Conrad et al., 2013). In spontaneous speech, speakers use "um" and "uh" to signal the start of expected minor or major delays (Clark & Fox Tree, 2002). Speakers can use these utterances to suggest they are searching for the right word, formulating an accurate description, preparing to repair what they just said, and deciding what to say next (Clark & Fox Tree, 2002). *Pauses* are periods of silences in speech and may signal that the speaker is

Direct correspondence to
Hanyu Sun, Westat, Rockville, MD, USA
E-mail: hanyusun@westat.com

organizing their thoughts or experiencing uncertainty (Fox Tree, 2002). Pauses are used for holding the floor or maintaining the turn (Maclay & Osgood, 1959). In replying to what the speaker has said, the addressee may start their reply with "um" or a pause, possibly indicating difficulty formulating a reply or that they are not comfortable with the topics under discussion (Fox Tree, 2010). *Repairs*—spontaneous corrections to an immediately prior utterance—occur when the speaker realizes they have made an error (Clark, 1994), for instance, "2 to 3 times a week... I mean frequently". In addition, speakers repair their speech to reduce ambiguity (Levelt, 1983). In everyday conversation, the speaker and the listener finely coordinate turn-taking (Sacks et al., 1974). *Overlapping* speech (i.e., where two speakers talk simultaneously) is often considered a breakdown in turn-taking. A conversation with large amount of overlapping speech often signals troubled communication (e.g., Leighton et al., 1971). Overall, the presence of disfluencies, pauses, repairs, and overlapping speech suggests that the speaker is experiencing communication difficulties to some extent.

Laughter alternates with speech to regulate the flow of the conversation (Jefferson, 1979). Not all laughter is an expression of amusement; laughter with high vowel resonances could reflect nervousness and submissiveness, while laughter with low vowel resonances could reflect coarseness and dominance (Kohler, 2008). Laughter by a first speaker may serve as an invitation to the addressee to laugh, resulting in shared laughter (Jefferson, 1979). Shared laughter has been linked to ending a topic, while declining to engage in reciprocal laughter rejects the proposed topic termination, leading to continued conversation on the topic (Holt, 2010). But laughter is not always shared. There are circumstances where it is inappropriate for the listener to accept the laughter invitation, for instance, when the speaker is talking about their troubles, the "troubles-teller" often produces an utterance and then laughs, but the "troubles-recipient" typically does not join in with the laughter (Jefferson, 1984). Laughter and smile can be considered as two extremes in the same continuum (Fried et al., 1988). The listener response smile seems to be equivalent to verbal acknowledgement like "mm-hmm" and "good", and is often accompanied by head nodding (Ekman, 1985). *Backchannels* (e.g., "uh huh" and "mmm") are thought to signal that the listener is attending to the speaker (Brunner 1979; Clark & Schaeffer, 1989) and possibly agrees with the speaker's message, consenting to the speaker's continuing to speak (Conrad et al., 2013).

Paralinguistic Behaviors in Standardized Interviews

Standardized interviews are different from everyday conversations. In standardized interviews, the interviewer directs the flow, and the interviewer-respondent interaction are expected to follow the paradigmatic question-answer sequence in which the interviewer asks a question and then the respondent selects a response option from those provided by the interviewer (Schaeffer and Maynard, 1996). For respondents, their sole task is to answer the interviewer's question. Interviewers, however, must keep the respondent engaged and ensure the collection of high-quality response data. To explore interaction in telephone and in-person survey interviews, researchers have studied the paralinguistic utterances of interviewers and respondents, especially disfluencies, hesitation, and pauses (e.g., Conrad et al., 2008; Conrad & Schober, 2021; Garbarski et al., 2016; Schober & Bloom, 2004; Schober et al., 2012). The presence of such behaviors can be used to flag respondent difficulty and, thus, increased likelihood of lower quality responses (e.g., Schaeffer & Dykema, 2011; Schober et al., 2012; Min et al., 2020). For instance, previous research has shown that respondents' disfluent answers—turns that include "um" and "uh", pauses, or repairs and restarts—are more likely to be unreliable than their fluent answers (Schober et al., 2012). Interviewers' disfluencies can affect survey outcomes by encouraging or discouraging sample member's decision to participate in a telephone interview (e.g., Conrad et al., 2013; Min et al., 2020). An interviewer who speaks without disfluencies may sound robotic to the respondent, giving the impression that they are only focused on completing the interviewing task rather than engaging with the respondent in a meaningful way. To the extent that this is detected by respondents, it can potentially reduce willingness to participate and if they do participate, their conscientiousness leading to reduced response quality. In addition, how the interviewer reacts to the respondent's "laughter invitation" may also affect the course of specific question-answer sequences (dialogue) and the overall interviewing experience. For instance, a telephone interviewer might use "smile voice" or quasi-laughter to create a sense of rapport with the respondent in the context of standardized interviewing (e.g., Lavin & Maynard, 2001). In addition, the more time interviewers spend smiling, the more likely the respondents were to provide an adequate response (Welles et al., 2022). (Adequate response is an indirect measure of data quality. A response is adequate if it matches one of the response categories provided in the survey question [Welles et al., 2022]).

Paralinguistic Behaviors in In-Person and Video Interviews

Video interviews may differ from in-person interviews in subtle ways due to the mediated nature of interaction. Because the interviewer and the respondent are in separate physical spaces, it is possible that they are less attuned to subtle social cues such as body language and nonverbal behaviors. At the same time, the increased sense of distance and privacy may make some respondents feel more comfortable sharing sensitive information. To the best of our knowledge, no research has been conducted comparing the prevalence of paralinguistic behaviors in in-person and video interviews. This paper explores that topic. In this study, we used the term "video interview" to refer to real-time, synchronized, two-way communication, mediated by video, distinguishing it from pre-recorded interviews where a video file of the interviewer reading the question is embedded in a web survey. Specifically, the current study asks the following research questions:

RQ1: *Does the prevalence of paralinguistic behaviors vary between in-person and video interviews?*

RQ2: *Does the prevalence of paralinguistic behaviors vary between in person and video interviews differently for interviewers versus respondents?*

RQ3: *Does data quality differ between modes in a corresponding way?*

Data and Method

To address these research questions, we examined the presence of paralinguistic behaviors in 15 in-person interviews and 12 video interviews in a laboratory experiment carried out in 2014. Five, professional female interviewers at a university research center administered standardized interviews in both modes (see Sun et al., 2020 for more information); respondents (14 females and 13 males, mean age of 46 [*SD* = 12], 24 of 27 held bachelor's degrees) were full-time staff employees of the University of Michigan, recruited via email and on-campus flyers, and randomly assigned to one of the two modes at the time of recruitment. Upon completion, they received a $15 token of thanks for their participation. The questions were on various topics including dietary behaviors, mental health, and sexual behaviors, adapted from existing national surveys such as the National Survey of Family Growth. The average interview length was about 35 minutes. The in-person interviews were conducted in an office, where the interviewer sat across a table from the respondent. The interviewer read the questions aloud from a laptop screen, and the respondent replied orally. The video interviews were conducted via Adobe Connect (www.adobe.com/products/adobeconnect.html). Both the interviewer and the respondent connected to the video call using the desktop computer. For the video respondents, most of their window was filled with the interviewer's video. The respondent's self-view video thumbnail appeared in the upper right corner. For the interviewer, most of their window was filled with the respondent's video. Interviewer's self-view video thumbnail appeared in the right upper corner. The interviewer shared their screen with the respondent to display showcards as needed.

The dialog between interviewers and respondents was audio-recorded in both in-person and video interviews. To transform the dialog into data that are amenable to quantitative analysis, we developed a coding scheme and then coded (in Sequence Viewer[1] [http://www.sequenceviewer.nl/]) the presence (Yes/No) of six paralinguistic behaviors for the interviewers and six for the respondents within each Q-A sequence (see Table 1); the behaviors included disfluencies, pauses, repairs, overlapping speech, laughter, and backchannels. In this study, disfluencies include fillers (e.g., “uh” and “um”). Laughter was coded as either interviewer-initiated or respondent-initiated during a Q-A sequence. Regarding overlapping speech, if the interviewer interrupted or cut off the respondent, we coded it as interviewer-initiated overlapping speech; if the respondent interrupted or cut off the interviewer’s speech, we coded it as respondent-initiated overlapping speech. We coded the presence of interviewer and respondent paralinguistic behaviors in each of 710 audio-recorded, question-answer (Q-A) sequences (i.e., for each question, we coded each conversational turn between the start of the question’s delivery [by the interviewer] and the start of the next question’s delivery). Two coders independently classified the interviewers’ and respondents’ behaviors; Cohen’s kappa = .75 was “substantial” (McHugh, 2012).

Table 1 Definition of the paralinguistic behaviors

Paralinguistic behaviors	Definition
Disfluencies	Fillers (e.g., “um” and “uh”)
Pauses	Periods of silence during a speaking turn
Repairs	Spontaneous corrections to an immediately prior utterance
Overlapping speech	Two speakers talk simultaneously
Laughter	Interviewer-initiated or respondent-initiated laughter during a Q-A sequence
Backchannels	Utterances such as, “uh huh” and “mmm”, to signal that the listener is attending to the speaker

Once the interactions were coded, we compared the presence of paralinguistic behaviors in the two modes. We first examined the presence of each paralinguistic behaviors by mode overall and separately for interviewer and respondent. We also computed dichotomous variables to measure communication difficulties overall and separately for interviewers and respondents. If a Q-A sequence included any of the paralinguistic behaviors—disfluencies, repairs, pauses, or overlapping speech—it is considered to exhibit communication difficulties. In the analysis, we accounted for the fact that interviewer and respondent utterances are clustered in an interview and used Wald chi-square tests for independence to compare the percentage of the presence of each paralinguistic behavior between the in-person and video interviews. We used PROC SURVEYFREQ in SAS 9.4 for the analysis and specified interviews (i.e., case id) as clusters.

1 We used Sequence Viewer 6.0d (http://www.sequenceviewer.nl/).

Because of how they were recruited, the participants comprised a convenience sample and thus the results are unweighted. We also adjusted for multiple comparisons using the Benjamini-Hochberg method (1995).

In addition, we fitted a crossed random effects logistic regression model to explore the effects of paralinguistic behaviors on the respondent's adequate response (Yes or No). A response is adequate if it matches one of the response categories provided in the survey question. As described earlier, the data includes 710 audio-recorded, question-answer (Q-A) sequences in total, combined across 15 in-person interviews and 12 video interviews conducted by 5 interviewers. The mean number of turns for each Q-A sequence for in-person and video interviews was 3.2 (SD = 1.8) and 3.6 (SD = 2.6), respectively. We initially fitted a two-level cross-classified random effects regression model with adequate response cross-classified by respondents and questions, nested within interviewers. However, the variance associated with interviewer was estimated to be zero in the unconditional model and therefore we simplified the statistical approach by fitting crossed random effects logistic regression models. The general model specification is as below:

$$\begin{aligned} &logit\,(\Pr(Y_{ij} = 1)) \\ &\quad = \beta_0 + \beta_1 \times Mode + \beta_2 \times Interviewer\ Paralinguistics \\ &\quad + \beta_3 \times Respondent\ Paralinguistics + u_i + v_j + \varepsilon_{ij} \\ &\quad u_i \sim N(0, \sigma_i^2),\ v_j \sim N(0, \sigma_j^2),\ \varepsilon_{ij} \sim N(0, \sigma^2) \end{aligned}$$

where $Y_{ij} = 1$ indicates adequate response for question i by respondent j, u_i represents the random effects associated with question i, v_j represents the random effects associated with respondent j, and ε_{ij} represents the residuals. The models are estimated using the default residual pseudo-likelihood (RSPL) in SAS 9.4 PROC MIXED with random intercepts for questions and respondents.

Results

RQ1 explores whether the prevalence of paralinguistic behaviors vary between in-person and video interviews. To address RQ1, we first examined the overall presence of paralinguistic behaviors by mode. As shown in Table 2, the overall prevalence of paralinguistic behaviors was at or below 30% in both in-person and video interviews. For instance, disfluencies occurred in 30.3% of speaking turns in in-person interviews and in 22.8% in video interviews, while laughter occurred in only of 2.8% of speaking turns in in-person and 13.8% in video interviews. Disfluencies and backchannels occurred significantly less frequently in video than in-person interviews while laughter and overlapping speech occurred significantly more often in video than in in-person interviews.

Table 2 The overall presence of paralinguistic behaviors by mode (by either the interviewer or the respondent)

Paralinguistic behaviors	In-person (%)		Video (%)		Wald chi-square test statistic[a]
	%	*n*[b]	%	*n*[b]	
Disfluencies	30.3	121	22.8	71	$F(1,26) = 1.68$, $p = .02$
Pauses	11.5	46	8.7	27	$F(1,26) = 0.50$, $p = .84$
Repairs	4.8	19	10.9	34	$F(1,26) = 9.79$, $p = .21$
Overlapping speech	10.5	42	15.8	49	$F(1,26) = 2.11$, $p = .01$
Laughter	2.8	11	13.8	43	$F(1,26) = 12.47$, $p = .01$
Backchannels	19.3	77	18.3	57	$F(1,26) = 0.03$, $p = .01$

Notes: Based on a total of 710 question-answer sequences irrespective of mode. The question-answer sequences were clustered by interview.

[a] Adjusted for multiple comparisons using the Benjamini-Hochberg method. The Wald Chi-square test statistics were from PROC SURVEYFREQ, specified interviews (i.e., case id) as clusters

[b] This column displays the number of times a paralinguistic behavior occurred.

RQ2 explores whether the prevalence of paralinguistic behaviors vary between in-person and video interviews differently for interviewers versus respondents. To address RQ2, we examined the presence of paralinguistic behaviors separately for interviewers and respondents. As shown in Figure 1a and 1b, there were significantly higher percentages of interviewer laughter ($F(1,26) = 12.93$, $p = .01$) and respondent laughter ($F(1,26) = 9.47$, $p = .03$) in video than in-person interviews. However, there were no significant differences between the two modes for other paralinguistic behaviors. As stated earlier, if the Q-A sequence included any of the paralinguistic behaviors—disfluencies, pauses, repairs, or overlapping speech—it is considered to exhibit communication difficulties. As shown in Figure 1a and 1b, interviewers experienced marginally more communication difficulties in video than in-person interviews ($F(1,26) = 3.30$, $p = .08$). Respondents experienced communication difficulties to the same extent in video and in-person interviews ($F(1,26) = 0.78$, $p = .38$).

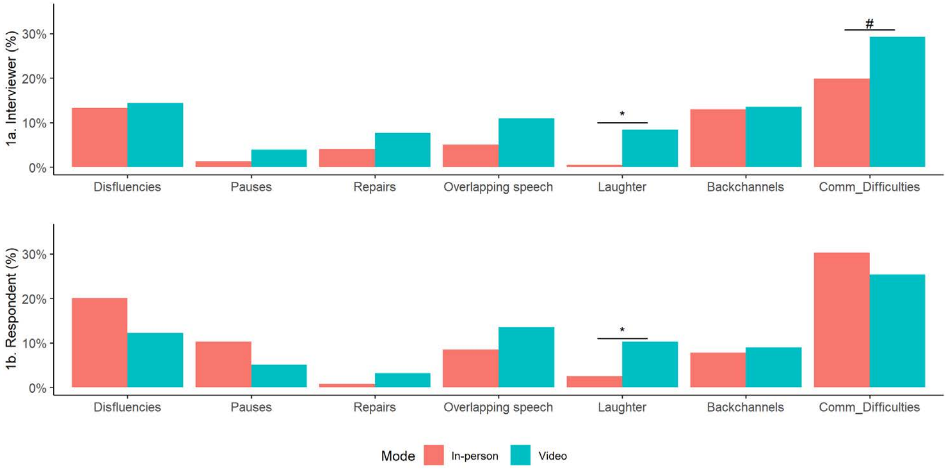

Figure 1 The presence of paralinguistic behavior by mode, separately for interviewers and respondents. Wald chi-square test statistics, specified interviews as clusters. Adjusted for multiple comparisons using the Benjamini-Hochberg method. * $p < .05$, # $p < .10$.

As noted earlier, laughter can serve different purposes in speech—it might be an expression of nervousness, or it could be someone laughing off their own troubles during a conversation. To better understand what drives the differences, we compared the prevalence of interviewer and respondent laughter when there were communication difficulties (either from the interviewer or the respondent) and when there were no difficulties for in-person and video interviews.

Figure 2 presents the prevalence of laughter by any communication difficulties in the Q-A sequence by mode. As shown in Figure 2a, no laugher was observed in the absence of communication difficulties in in-person interviews. It was only when there were communication difficulties that there was interviewer laughter and respondent laughter in in-person interviews. The percentage of respondent laughter was higher than the percentage of interviewer laughter in in-person interviews. As shown in Figure 2b, the percentage of interviewer laughter and the percentage of respondent laughter were higher when there were communication difficulties as compared to when there were no difficulties in video interviews. The prevalence of laughter was much lower in in-person than video interviews. For instance, as shown in Figure 2a, 6.1% of the Q-A sequences contain respondent laughter where there were no communication difficulties in video interviews as compared to 0% of respondent laughter when there were no communication difficulties in in-person interviews (a condition that was never actually observed as all sequences in this mode contained communication difficulty either from the interviewer or the respondent). The percentage of respondent laughter increased to 16.8% when there were communication difficulties in video interviews as shown in Figure 2b. The findings are consistent when separately examining the interviewer and respondent communication difficulties (see Figures A1 and A2 in the Appendix).

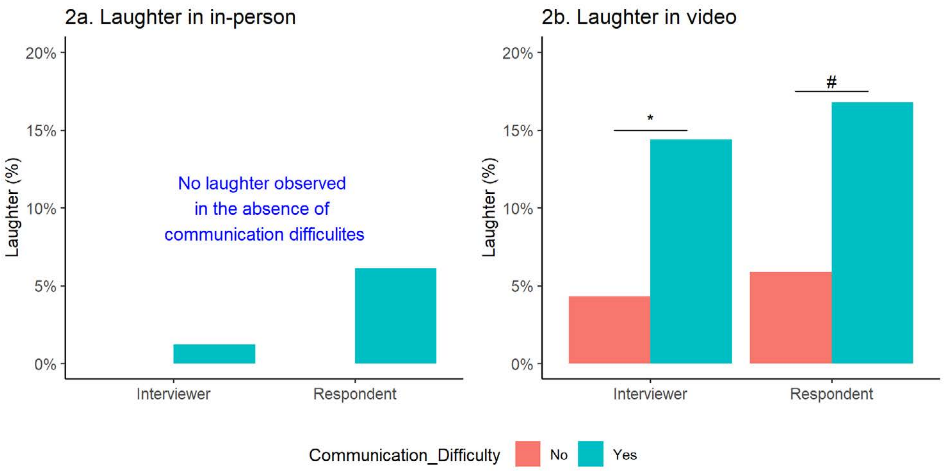

Figure 2 The prevalence of laughter by communication difficulties (by either the interviewer or the respondent) by mode. Adjusted for multiple comparisons using the Benjamini-Hochberg method. Chi-square tests cannot be computed if at least one table cell has 0 frequency. * $p < .05$, # $p < .10$.

RQ3 explores whether the presence of paralinguistic behaviors affect data quality by mode. To address RQ3, we fitted a crossed random effects logistic regression model to predict adequate response—an indirect measure of data quality—by the paralinguistic behaviors (present or absent) and mode (in-person or video). As Shown in Table 3, we found that respondents were less likely to provide an adequate response when interviewers repaired their speech (e.g., “How long has it been since you last used any prescription transit uh tranquilizer…”), when respondents were disfluent, and when respondents interrupted the interviewer, after controlling for the other paralinguistic behaviors and mode. But there were no significant two-way interactions between any of the paralinguistic behaviors and mode on the prevalence of adequate responses (results not shown). As a result, although we found significant differences between laughter in in-person and video interviews, this did not differentially impact the prevalence of respondents’ adequate responses (i.e., was not related to [our measure of] data quality).

Table 3 Crossed random effects logistic regressions predicting adequate response by the presence of paralinguistic behaviors and mode

Parameters	*b (SE)*	
Intercept	3.58***	(0.36)
Video interview (ref: In-person)	−0.09	(0.39)
Interviewer paralinguistic behaviors		
Disfluencies (ref: Not present)	0.72	(0.54)
Repairs (ref: Not present)	−1.10#	(0.59)
Pauses (ref: Not present)	1.14	(1.14)
Overlapping speech (ref: Not present)	0.74	(0.65)
Laughter (ref: Not present)	0.19	(0.87)
Backchannels (ref: Not present)	−0.21	(0.43)
Respondent paralinguistic behaviors		
Disfluencies (ref: Not present)	−1.60***	(0.38)
Repairs (ref: Not present)	−1.26	(0.87)
Pauses (ref: Not present)	0.06	(0.53)
Overlapping speech (ref: Not present)	−1.47**	(0.48)
Laughter (ref: Not present)	−0.37	(0.59)
Backchannels (ref: Not present)	−0.29	(0.53)
$\sigma^2_{int:question}$	0.89	(0.43)
$\sigma^2_{int:respondent}$	0.22	(0.27)
n	710	

Notes: ref = reference. *** $p < .001$, ** $p < .01$, * $p < .05$, # $p < .10$.

Discussion and Conclusion

In this study, we compared the presence of interviewer and respondent paralinguistic behaviors between in-person and video interviews. One paralinguistic behavior differed between modes: Interviewers and respondents laughed significantly more in video than in in-person interviews. Although we did not observe any mode differences for the other paralinguistic behaviors, the presence of laughter was correlated with the presence of communication difficulties. One interpretation of this finding is that laughter was functioning as a “social lubricant” to keep the conversation flowing when it was disrupted by these behaviors.

The study was conducted in 2014, a time when video as a mode of data collection was new to interviewers and respondents. While it is possible that technical issues such as bandwidth limitation, lags in transmission, and poor audio/video quality may have led to communication difficulties, we are not aware of any technical problems having occurred during the interviews. It is plausible that video requires participants to split their attention between planning what to say and monitoring activities across multiple views (i.e., the interviewer needed to monitor the respondent's video and survey instrument, the respondent may have monitored their self-view video thumbnail), therefore increasing the prevalence of paralinguistic behaviors that indicate communication difficulties. The one mode effect on paralinguistic behavior (i.e., laugher) observed in the current study was small but persistent when the data were collected in 2014, and video was relatively new. This effect of mode on paralinguistic behavior is likely to be even smaller now that video is common in everyday communication. Future research may explore this by conducting similar studies to test mode differences between in-person and video interviews, since video technologies are now widespread and familiar to much of the public. In a more contemporary test, the researcher would be wise to recruit a larger sample of participants from a more general population.

In addition, we found that the presence of interviewer repairs, respondent disfluencies, and respondent-initiated overlapping speech affected adequate response. However, there were no significant two-way interactions between any of the paralinguistic behaviors and mode on adequate response. The very similar patterns of interaction in video and in-person interviews are consistent with several findings in the literature indicating very similar respondent experiences and data quality in video and in-person interviews. As previously indicated, Sun et al. (2020) reported similar levels of respondents' rated rapport with the interviewer and Endres et al. (2022) reported similar quality data in the two modes, in particular equivalent length of open responses, straightlining, item nonresponse, and social desirability bias.

The study has some limitations. When the study was conducted (2014), video was still a relatively new method of data collection for both interviewers and respondents. The communication difficulties caused by technical issues have likely been greatly reduced today due to improved Internet connectivity (at least in urban areas) and widespread familiarity with video conferencing platforms. It was difficult to assess the impact of technical issues on communication difficulties arising from the mode difference between in-person and video interviews because we only had audio recordings of the interviews. Technical issues were a known aspect of video interviews in 2014. However, recent studies have found that respondent experiences and data quality are similar between video and in-person interviews (e.g., Endres et al., 2022). This has led us to believe that part of the laughter we observed in 2014 may have been due to technical issues more than mode differences. Future research should tease apart technical issues from communication difficulties in testing mode effects on paralinguistic behaviors. For instance, developing a coding scheme to identify technical issues and then comparing Q-A sequences with and without these difficulties could help answer how paralinguistic behaviors vary by mode. To the extent that communication difficulties appear when there are no technical difficulties, this might suggest benefit in training interviewers on how to best multitask while maintaining respondent engagement in video interviews.

Nevertheless, these findings strongly argue for using video to conduct survey interviews when the measurement goals of the research require in-person data collection. Although we observed slightly different interaction patterns in video than in-person interviews, the adequacy of the answers in the two modes was indistinguishable.

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Brown, G. (1977). *Listening to spoken English* (2nd ed.). Routledge. https://doi.org/10.4324/9781315538518

Brunner, L. J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology, 37*(5), 728–734. https://doi.org/10.1037/0022-3514.37.5.728

Clark, H. H. (1994). Managing problems in speaking. *Speech Communication, 15*(3–4), 243–250. https://doi.org/10.1016/0167-6393(94)90075-2

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*(2), 259–294. https://doi.org/10.1207/s15516709cog1302_7

Conrad, F. G., Broome, J. S., Benkí, J. R., Kreuter, F., Groves, R. M., Vannette, D., & McClain, C. (2013). Interviewer speech and the success of survey invitations. *Journal of the Royal Statistical Society Series A: Statistics in Society, 176*(1), 191–210. https://doi.org/10.1111/j.1467-985X.2012.01064.x

Conrad, F. G., & Schober, M. F. (2021). Clarifying question meaning in standardized interviews can improve data quality even though wording may change: A review of the evidence. *International Journal of Social Research Methodology, 24*(2), 203–226. https://doi.org/10.1080/13645579.2020.1824627

Conrad, F. G., Schober, M., & Dijkstra, W. (2008). Cues of communication difficulty in telephone interviews. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W., Link, & R. L. Sangster (Eds.), *Advances in telephone survey methodology* (pp. 212–230). John Wiley & Sons. https://doi.org/10.1002/9780470173404.ch10

Eckman, P. (1985). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. Norton.

Endres, K., Hillygus, D. S., DeBell, M., & Iyengar, S. (2022). A randomized experiment evaluating survey mode effects for video interviewing. *Political Science Research and Methods, 11*(1), 144–159. https://doi.org/10.1017/psrm.2022.30

Fried, I., Wilson, C. L., MacDonald, K. A., & Behnke, E. J. (1998). Electric current stimulates laughter. *Nature, 391*, 650. https://doi.org/10.1038/35536

Fox Tree, J. E. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes, 34*(1), 37–55. https://doi.org/10.1207/S15326950DP3401_2

Garbarski, D., Schaeffer, N. C., & Dykema, J. (2016). Interviewing practices, conversational practices, and rapport: Responsiveness and engagement in the standardized survey interview. *Sociological Methodology, 46(*1), 1–38. https://doi.org/10.1177/0081175016637890

Hancock, J. T. (2004). Verbal irony use in face-to-face and computer-mediated conversations. *Journal of Language and Social Psychology, 23*(4), 447–463. https://doi.org/10.1177/0261927X04269587

Holt, E. (2010). The last laugh: Shared laughter and topic termination. *Journal of Pragmatics, 42*(6), 1513–1525. https://doi.org/10.1016/j.pragma.2010.01.011

Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 79–96). Irvington.

Kohler, K. J. (2008). "Speech-smile", "speech-laugh", "laughter" and their sequencing in dialogic interaction. *Phonetica*, *65*(1–2), 1–18. https://doi.org/10.1159/000130013

Kokosak A., Sun, H., & Yan, T. (2024, May 15–17). *How do behavior coding results differ by mode in a multimode survey?* [Paper presentation]. 79th Annual Conference of the American Association for Public Opinion Research (AAPOR), Atlanta, GA, United States. https://aapor.confex.com/aapor/2024/meetingapp.cgi/Paper/2649

Lavin, D., & Maynard, D. W. (2001). Standardization vs. rapport: Respondent laughter and interviewer reaction during telephone surveys. *American Sociological Review, 66*(3), 453–479. https://doi.org/10.2307/3088888

Leighton, L. A., Stollak, G. E., & Ferguson, L. R. (1971). Patterns of communication in normal and clinic families. *Journal of Consulting and Clinical Psychology, 36*(2), 252–256. https://doi.org/10.1037/h0030759

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41–104. https://doi.org/10.1016/0010-0277(83)90026-4

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15*(1), 19–44. https://doi.org/10.1080/00437956.1959.11659682

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282. https://biochemia-medica.com/en/journal/22/3/10.11613/BM.2012.031

Min, B. H., Schaeffer, N. C., Garbarski, D., & Dykema, J. (2020). The action structure of recruitment calls and its analytic implications: The case of disfluencies. *Journal of Official Statistics, 36*(3), 529-559. https://doi.org/10.2478/jos-2020-0028

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696–735. https://doi.org/10.2307/412243

Schober, M. F., & Bloom, J. E. (2004). Discourse cues that respondents have misunderstood survey questions. *Discourse processes, 38*(3), 287–308. https://doi.org/10.1207/s15326950dp3803_1

Schober, M. F., Conrad, F. G., Dijkstra, W., & Ongena, Y. P. (2012). Disfluencies and gaze aversion in unreliable responses to survey questions. *Journal of Official Statistics, 28*(4), 555–582. https://www.proquest.com/scholarly-journals/disfluencies-gaze-aversion-unreliable-responses/docview/1266766372/se-2?accountid=16283

Sun, H., Conrad, F. G., & Kreuter, F. (2020). The relationship between interviewer-respondent rapport and data quality. *Journal of Survey Statistics and Methodology, 9*(3), 429–448. https://doi.org/10.1093/jssam/smz043

Foucault Welles, B., Sun, H., & Miller, P. V. (2022). Nonverbal behavior in face-to-face survey interviews: An analysis of interviewer behavior and adequate responding. *Field Methods, 34*(1), 52–68. https://doi.org/10.1177/1525822X221075305

Appendix

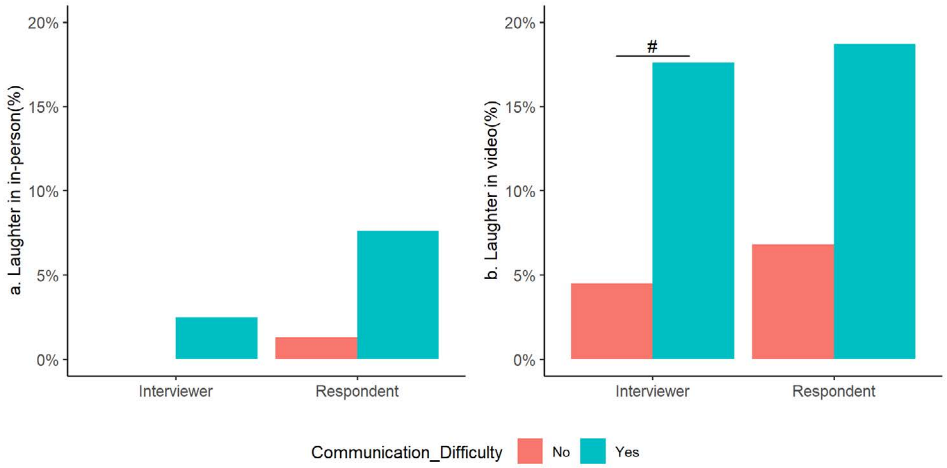

Figure A1 The prevalence of laughter by interviewer communication difficulties by mode. Adjusted for multiple comparisons using the Benjamini-Hochberg method. Chi-square tests cannot be computed if at least one table cell has 0 frequency. # $p < .10$.

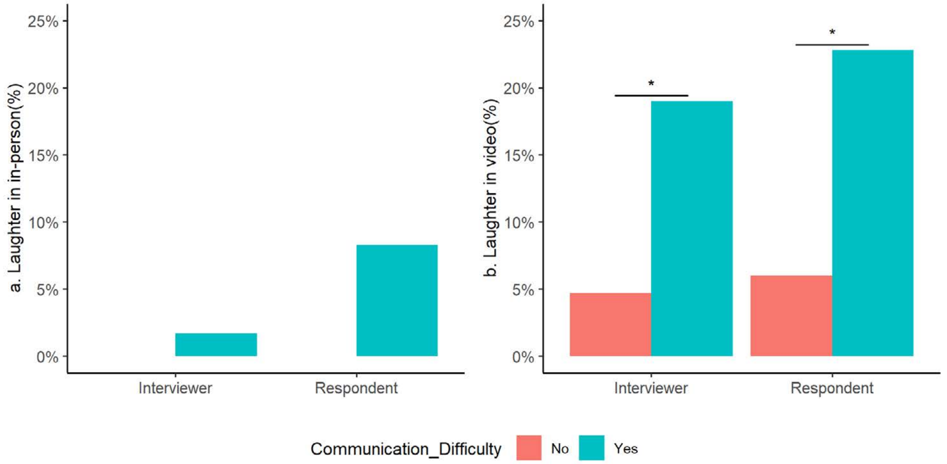

Figure A2 The prevalence of laughter by respondent communication difficulties by mode. Adjusted for multiple comparisons using the Benjamini-Hochberg method. * $p < .05$.

Online First publication: 2025-11-21