Fall 2018

# Why Do Cell Phone Interviews Last Longer? A Behavior Coding Perspective

Jerry Timbrook

Kristen Olson

Jolene Smyth

# Why Do Cell Phone Interviews Last Longer? A Behavior Coding Perspective

Jerry Timbrook, Kristen Olson, and Jolene D. Smyth

University of Nebraska–Lincoln

*Corresponding author* — Jerry Timbrook, University of Nebraska–Lincoln, Department of Sociology, 711 Oldfather Hall, Lincoln, NE 68588-0324, USA; *email* jtimbrook2@unl.edu

## Abstract

Why do telephone interviews last longer on cell phones than landline phones? Common explanations for this phenomenon include differential selection into subsets of questions, activities outside the question-answer sequence (such as collecting contact information for cell-minute reimbursement), respondent characteristics, behaviors indicating disruption to respondents' perception and comprehension, and behaviors indicating interviewer reactions to disruption. We find that the time difference persists

Published in *Public Opinion Quarterly* 82:3 (Fall 2018), pp 553–582.
doi 10.1093/poq/nfy022

Published August 14, 2018.
Jerry Timbrook is a graduate student in the Department of Sociology, University of Nebraska– Lincoln, Lincoln, NE, USA. Kristen Olson is the Leland J. and Dorothy H. Olson Associate Professor and vice chair in the Department of Sociology at the University of Nebraska–Lincoln, Lincoln, NE, USA. Jolene D. Smyth is an associate professor in the Department of Sociology and director of the Bureau of Sociological Research at the University of Nebraska–Lincoln, Lincoln, NE, USA.

This material is based upon work supported by the National Science Foundation [Grant No. SES-1132015 to K.O.]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

An earlier version of this paper was presented at the Annual Meeting of the American Association for Public Opinion Research, May 2016, Austin, TX, USA.

even when we focus only on the question-answer portion of the interview and only on shared questions (i.e., eliminating the first two explanations above). To learn why the difference persists, we use behavior codes from the U.S./Japan Newspaper Opinion Poll, a dual-frame telephone survey of US adults, to examine indicators of satisficing, line-quality issues, and distraction. Overall, we find that respondents on cell phones are more disrupted, and that the difference in interview duration occurs because cell phone respondents take longer to provide acceptable answers. Interviewers also slow their speed of speech when asking questions. A slower speaking rate from both actors results in a longer and more expensive interview when respondents use cell phones.

## Introduction

In June 2017, 52.0 percent of US adults lived in households that were reachable only on a cell phone, with an additional 16.3 percent using cell phones as their primary device despite also having landline service (Blumberg and Luke 2017). Including cell phone respondents in random-digit-dial telephone surveys thus has evolved from useful to necessary. The Pew Research Center announced that 75 percent of its telephone survey respondents will come from cellular numbers (McGeeney 2016), and other researchers have suggested completely abandoning landline telephone interviews (Peytchev and Neely 2013; Gundersen et al. 2014; Kennedy, McGeeney, and Keeter 2016).

Despite the utility of dual-frame samples in improving coverage, researchers have expressed concern that the quality of responses may differ between these two devices (Lavrakas et al. 2007). Yet, studies find few differences on indicators of data quality between cell phone and landline surveys (Brick et al. 2007; Witt, ZuWallack, and Conrey 2009; Lavrakas, Tompson, et al. 2010; Kennedy 2010; Kennedy and Everett 2011; Lynn and Kaminska 2012).

One consistent exception is that cell phone surveys last longer than landline surveys (Kuusela and Notkola 1999; Keeter and Kennedy 2006; Brick et al. 2007; Vicente, Reis, and Santos 2009; Lynn and Kaminska 2012). Because cell phone surveys are more expensive than landline surveys (McGeeney 2016), understanding why this difference occurs is important. Although multiple hypotheses exist for this difference, no study has empirically established why there is a

difference in length between landline and cell phone surveys. This paper expands previous work to examine features of the interviewer-respondent interaction during the interview. In particular, we address the following four questions:

1. Does the difference in interview length between landline and cell phone interviews persist when looking only at the question-answer portion of the interview (i.e., setting aside activities that occur before or after the questions) and for the same set of questions?
2. Do respondent characteristics account for the difference in length?
3. Are there differences in the rates of occurrence of interview behaviors between landline and cell phone interviews, and do these differences account for the difference in interview length?
4. Do the same interviewer and respondent behaviors (e.g., asking questions or providing an answer) take different amounts of time in landline and cell phone interviews and account for the difference in length between cell phone and landline interviews?

## Background

Interviews on cell phones take up to about 15 percent longer to complete than interviews on landline phones (**Table 1**). This effect has been replicated in the United States (Keeter and Kennedy 2006; Brick et al. 2007), Finland (Kuusela and Notkola 1999), Portugal (Vicente, Reis, and Santos 2009), and Hungary (Lynn and Kaminska 2012).

As cell phone interviews cost 1.5 to 2 times that of landline interviews (McGeeney 2016), survey researchers have a financial incentive to understand this difference in interview length. If cell phone interviews last an average of 1 minute longer than their landline counterparts, a telephone survey with 500 cell phone interviews will require 500 more interviewer minutes than the same survey with landline interviews. This difference is equivalent to almost 42 completed 12-minute interviews. We discuss three potential causes of this difference below.

**Table 1.** Reported interview length by device for five studies

| | *Mean length (in minutes)* | | | | |
|---|---|---|---|---|---|
| *Study* | *Landline* | *Cell* | *Diff.* | *% Cell exceeds landline* | *Topic of study* |
| Kuusela & Notkola 1999 | 4.0 | 4.3 | 0.3 | 7.5% | Labor |
| Keeter & Kennedy 2006 | 10.2 | 11.8 | 1.6 | 15.7% | Politics |
| Brick et al. 2007 | 8.2 | 8.9 | 0.7 | 8.5% | Phone use and social issues |
| Vicente et al. 2009 | 10.9 | 12.0 | 1.1 | 10.1% | Internet use |
| Lynn & Kaminska 2012a | 14.6 | 16.7 | 2.1 | 14.4% | Social issues |

a. For this study, mean length by device was provided by the authors.

### *Nonsurvey Activities and Skip Patterns*

Activities unrelated to the substantive survey questions may lead to the difference in interview length. Two such activities include finding a private place to answer (Lynn and Kaminska 2012) or collecting contact information for incentive/reimbursement delivery (Keeter and Kennedy 2006). Additionally, cell phone respondents may be predisposed to select into certain skip patterns, yielding additional survey questions (Kuusela and Notkola 1999).

These hypotheses posit that cell phone interviews are longer due to interactions that occur outside the question-answer sequence. Yet, both cellular phones and cordless landline phones allow respondents to find a private place to answer questions. Additionally, few cell respondents complete surveys in public (Brick et al. 2007). Although post-survey collection of information for incentives for cell phone interviews could increase length, this time could be offset by within-household selection procedures for landline interviews that are not typically used for cell phones (typically considered to be personal, not household, devices) (Kennedy 2010). Thus, the difference in interview length should persist when examining the same set of questions for both landline and cell respondents, excluding these non-survey activities.

### *Respondent Characteristics*

Socio-demographic differences may account for the disparity in length between devices (Nathan 2001). Older respondents have longer

interviews than younger respondents (Fricker et al. 2005; Couper and Kreuter 2013; Olson and Smyth 2015), perhaps because they have decreased working memory capacity (Salthouse 1991), more difficulty hearing (Linville 2001), or fewer time demands (U.S. Bureau of Labor Statistics 2016). However, landline users are typically older than cell phone users (Blumberg and Luke 2017). Thus, if the difference were due to age, we would expect landline interviews to last longer. Similarly, education is positively correlated with cognitive skill (Ceci 1991); respondents with lower education levels may take longer to process survey questions. Yet, level of education in the population varies only slightly across devices (Blumberg and Luke 2017). Finally, men speak more quickly than women (Verhoeven, De Pauw, and Kloots 2004). In the general population, men and women are equally likely to be cell phone users only (Blumberg and Luke 2017), but differential nonresponse and selection across frames may lead to differences in gender composition in respondent pools. As older respondents are more likely to be on landlines, and as education and sex are not expected to vary across device types, respondent characteristics should not account for the increased length of cell phone interviews.

### *Disruption Of Perception And Comprehension*

The first two steps of answering a survey question are perception and comprehension (Jenkins and Dillman 1997; Tourangeau, Rips, and Rasinski 2000). For perception, if a respondent is unable to hear a survey question, extra effort may be required to negotiate understanding. Similarly, comprehension difficulties may take time to resolve.

Disruptions to perception and comprehension may be more pronounced during cell interviews due to reduced sound quality on cell phones (Lavrakas, Blumberg, et al. 2010; Lavrakas, Tompson, et al. 2010; Kennedy and Everett 2011). Respondents and interviewers may have difficulty hearing one another on a cell phone, requiring the actors to repeat themselves. However, literature comparing audio fidelity or requests for repeating information across landlines and cell phones is sparse.

Distractions such as background noise (Schwarz et al. 1991; Lavrakas, Blumberg, et al. 2010) and multitasking (Holbrook, Green, and Krosnick 2003; Schober et al. 2015) may be more prevalent during

cell phone interviews (Lavrakas, Blumberg, et al. 2010), explaining the difference in survey length. These distractions may divide a respondent's attention, leaving them with fewer cognitive resources to perceive and comprehend survey questions (Lynn and Kaminska 2012). However, interviewers rate cell and landline respondents as equally distracted, and both interviewer ratings and respondent self-reports indicate no difference in multitasking across devices (Keeter and Kennedy 2006; Kennedy 2010; Lynn and Kaminska 2012). Additionally, landline respondents are more likely than cell respondents to pay attention to instructions on longer questions, but not on shorter questions (Kennedy and Everett 2011).

Respondents who are distracted may also engage in satisficing behaviors to reduce cognitive effort (Krosnick 1991), prompting more follow-up from the interviewer to obtain an adequate answer. But no differences have been found across cell phone and landline interviews on indicators of satisficing such as straightlining, length of open-ended responses, item nonresponse rates, or response-order effects (Brick et al. 2007; Witt, ZuWallack, and Conrey 2009; Kennedy 2010; Lavrakas, Blumberg, et al. 2010; Kennedy and Everett 2011; Lynn and Kaminska 2012).

Previous literature thus finds few, if any, differences in disruption to perception and comprehension between cell and landline telephone interviews. However, interviewer ratings, self-reports of multitasking or distraction, and data-quality indicators derived from responses themselves may not be sensitive enough to detect differences in cognitive states between respondents using the two devices (Fowler and Cannell 1996). Respondent and interviewer behaviors during an interview may serve as better indicators of difficulties at the perception and comprehension stages. For example, disrupted respondents may take more conversational turns to provide acceptable answers, and may provide unacceptable answers more often. Perception and comprehension difficulty can also be signaled by conversational behaviors such as disfluencies (e.g., "uh," "um") (Schober and Bloom 2004), asking "What did you say?" (Yont, Hewitt, and Miccio 2002), or asking for a question to be repeated (Miller et al. 2011; Thrasher et al. 2011). Slower speech can indicate that a speaker is coping with a high cognitive load and distraction (Berthold and Jameson 1999). Cell respondents who are distracted may need more time to construct responses

that are equal in quality to those of landline respondents. Issues with audio quality might be signaled when actors make explicit comments about poor telephone service during an interview, when unintelligible audio is identified in interview recordings, or when actors interrupt each other often (Hammer and Reichl 2005).

If interviewers notice respondents engaging in these behaviors, they may change their own question-asking behavior accordingly. Interviewers may probe more often if they notice that respondents are disrupted, offer more clarifications, use positive feedback (i.e., "We really appreciate your answers") more frequently to motivate respondents, verify answers more often, or speak slower to facilitate understanding. This perspective follows communication accommodation theory, which posits that when individuals interact they use context clues from their exchange to converge on a style of communication that is effective for both actors, such as by talking more quickly or slowly to match their conversational counterpart (Giles, Coupland, and Coupland 1991). Additionally, interviewers contending with poor sound quality may experience increased cognitive burden, thus increasing interviewer disfluencies.

There are two ways that these behaviors may affect interview length. First, if cell phone interviews have more of these behaviors, this may lead to a longer call. Second, some behaviors may take longer on a cell phone than a landline. For example, Schober et al. (2015) find that conversational turns last longer during text messaging interviews than during telephone voice interviews, but that text messaging interviews had fewer conversational turns. Therefore, respondents and interviewers on cell phones may not necessarily take more turns to complete the question-answer sequence, but they may take longer during those turns. To our knowledge, no study has empirically examined differences in the rates of occurrence or duration of individual behaviors between landline and cell interviews, nor whether the same interview behaviors contribute differentially to interview length across devices. In this paper, we empirically examine whether the difference in interview length across devices persists within the question-answer portion of the survey and whether the difference can be explained by (1) respondent characteristics or by differences in (2) the rates of occurrence or (3) duration of interview behaviors across devices.

## Data and Methods

The data come from the U.S./Japan Newspaper Opinion Poll (NOP), a dual-frame random-digit-dial telephone survey of US adults fielded by Gallup on November 18–25, 2013 (AAPOR RR1 = 7.4 percent). The NOP consisted of 57 closed-ended political opinion items, and 15 demographic items (Online Appendix A). The Bureau of Sociological Research at the University of Nebraska–Lincoln transcribed a stratified random subset of 438 interviews conducted in English. Four respondents did not indicate the type of device on which they were responding, and five partial interviews are removed from analysis, resulting in $n$ = 429 interviews (245 landline, 184 cell) conducted by 31 interviewers.

The dependent variable is the number of minutes spent in the asking and answering of questions during the survey, excluding time spent on recruitment, consent, within-household selection, and anything else occurring outside a question-answer sequence. Interview length was calculated by summing the length of each conversational turn1 in deciseconds (the unit used by the software Sequence Viewer; Dijkstra 1999) from the first question asked to the final question asked in the survey. The time associated with one question asked only of a portion of landline respondents was removed from the total interview length. Thus, the dependent variable is the total amount of time (transformed to minutes) for the entire question-answer portion of the interview on identical sets of questions for landline and cell phone respondents. The interview averaged 12.30 minutes. We use a natural log transformation to account for the skewed distribution of interview length in our models.

Our independent variable of interest is the self-reported device on which respondents were interviewed (landline = 0, cell = 1). Overall, 57 percent responded on a landline telephone and 43 percent responded by cell phone.

Our key independent variables explaining the device effect start with respondent characteristics that may affect interview duration:

1. Conversational turns were operationalized as each period of uninterrupted speech by a single actor, with turns ending when an actor finished speaking or was interrupted by another actor. Instances of overlapping speech were counted as a single turn.

**Table 2.** Descriptive statistics for interview length, respondent controls, and respondent characteristics, overall and by device

| | *Overall (n = 429)* | *Landline (n = 245)* | *Cell phone (n = 184)* | *Diff.* |
|---|---|---|---|---|
| Total interview length (in mins) | 12.30 | 11.80 | 12.97 | −1.17** |
| *Respondent controls* | | | | |
| Race | | | | |
| White, not Hispanic (ref) | 82.98% | 84.90% | 80.43% | 4.47 |
| All other races | 17.02% | 15.10% | 19.57% | |
| Yearly income | | | | |
| Less than $55,000 (ref) | 38.46% | 40.41% | 35.87% | 4.54 |
| $55,000 or more | 61.54% | 59.59% | 64.13% | |
| *Respondent characteristics* | | | | |
| Age (in years) | 54.82 | 60.88 | 46.88 | 14.00* |
| Education | | | | |
| High school graduate or less (ref) | 26.11% | 26.12% | 26.09% | 0.03 |
| Some college | 21.91% | 23.67% | 19.57% | 4.10 |
| College graduate or more | 51.98% | 50.20% | 54.35% | −4.15 |
| Sex | | | | |
| Male (ref) | 57.34% | 47.35% | 70.65% | −23.30*** |
| Female | 42.66% | 52.65% | 29.35% | |

* $p < 0.05$ ; ** $p < 0.01$ ; *** $p < 0.001$

age (in years), education (three dichotomous variables: high school or less, some college, college graduate or higher), and interviewer-coded respondent sex (**Table 2**). Missing, don't know, and refusal answers to these questions were imputed to each variable's mean or modal category (≤ 1.6 percent missing). We also control for interviewer age as a continuous variable (mean = 34), sex (58.06 percent female), and tenure in years of interviewing experience (mean = 1.83 years); all 31 interviewers were white. To account for compositional differences across devices (Blumberg and Luke 2017), we include respondent income and race.

Our next set of measures proxy for satisficing and difficulty with perception and comprehension. These come from behavior codes of the interview itself. Behavior coding is a method of systematically coding the interaction between respondents and interviewers during survey completion (Fowler and Cannell 1996; Schaeffer and Maynard 1996; Ongena and Dijkstra 2006). Behavior codes allow for a detailed examination of respondent cognitive states during an interview and

provide descriptions of what happened at a specific moment during a survey (Fowler and Cannell 1996). As such, they provide measures less prone to recall error than post-survey interviewer evaluations or respondent self-reports.

Trained undergraduates behavior-coded each conversational turn of each interview using the Sequence Viewer software program (Dijkstra 1999). We use codes for the actor (interviewer or respondent), initial action (e.g., the interviewer asked a question), assessment of the initial action (e.g., the interviewer read the question with changes), whether the actor uttered a disfluency on this turn, and whether one actor interrupted another on this turn. Master coders coded a random subset of 10 percent of the coded transcripts to evaluate inter-coder reliability. Kappa values for all but two behaviors exceed 0.60 (moderate agreement) (Landis and Koch 1977; McHugh 2012). These two behaviors were rare (< 2 percent of all behaviors); percent agreement (an alternative to kappa for rare events; see Viera and Garrett [2005]) exceeds 60 percent for both behaviors.

Using these behavior codes, we calculated indicators of satisficing and disruption to perception and comprehension. Because interview length may differ due to the rates of occurrence *or* the duration of these behaviors, we calculated our indicators across all turns for each call in two ways: (1) the total number of conversational turns on which each behavior occurs (for parsimony, we refer to these as "frequencies" of these behaviors); and (2) the total length (in minutes) for each of these behaviors (duration).

### *Frequencies of Interview Behaviors*

We start by examining behavioral indicators of satisficing (**Table 3**) to explore the role that cognitive shortcutting plays in interview length. First, we create an indicator of acquiescence (Krosnick 1991) as the total number of turns on which a respondent agrees with the interviewer. We then count the number of turns on which a respondent provides a "don't know" answer or refuses to answer a survey question (Krosnick et al. 2002). Finally, we examine the number of turns a respondent or interviewer makes a time-related comment about the survey (i.e., asking how much longer is left in the interview).

The next set of independent variables proxy for disruption to perception and comprehension (see **Table 3**). The first proxy is the total

**Table 3.** Summary of behaviors indicating satisficing and disruption to perception or comprehension

| | *Overall mean* | *(SE)* | *Landline mean* | *(SE)* | *Cell phone mean* | *(SE)* | *LL – Cell diff.* |
|---|---|---|---|---|---|---|---|
| Total number of turns | 203.82 | (4.09) | 206.20 | (4.51) | 200.65 | (5.61) | 5.55 |
| ***Behaviors indicating satisficing*** | | | | | | | |
| Number of turns respondent agrees with interviewer | 0.21 | (0.06) | 0.27 | (0.08) | 0.13 | (0.04) | 0.14# |
| Number of turns respondent gives DK/REF answers | 2.15 | (0.29) | 2.55 | (0.40) | 1.61 | (0.34) | 0.94# |
| Number of turns respondent comments about duration | 0.08 | (0.02) | 0.07 | (0.02) | 0.09 | (0.02) | –0.02 |
| Number of turns interviewer comments about duration | 0.47 | (0.04) | 0.44 | (0.04) | 0.50 | (0.06) | –0.06 |
| ***Respondent indicators of disruption to perception or comprehension*** | | | | | | | |
| Number of turns with an acceptable answer | 65.13 | (0.70) | 65.11 | (0.80) | 65.14 | (1.04) | –0.03 |
| Number of turns with an unacceptable answer | 7.45 | (0.44) | 7.82 | (0.54) | 6.95 | (0.71) | 0.87 |
| Number of turns with respondent disfluencies | 17.36 | (0.59) | 16.20 | (0.82) | 18.90 | (1.07) | –2.70# |
| Number of turns respondent says "What?" | 0.75 | (0.08) | 0.76 | (0.10) | 0.74 | (0.11) | 0.02 |
| Number of turns respondent asks interviewer to repeat question, definition, or response options | 2.72 | (0.20) | 2.85 | (0.28) | 2.54 | (0.20) | 0.31 |
| Number of turns respondent interrupts interviewer | 6.01 | (0.32) | 6.37 | (0.41) | 5.52 | (0.42) | 0.85 |
| Respondent speed of speech (in words per minute) | 82.46 | (1.79) | 90.73 | (1.78) | 71.46 | (1.59) | 19.27*** |
| Percent of interviews with comments about line quality | 6.76% | (1.34%) | 5.71% | (1.34%) | 8.15% | (1.81%) | –2.44% |
| Percent of interviews with unintelligible respondent audio | 50.82% | (2.99%) | 46.12% | (3.41%) | 57.07% | (4.15%) | –10.95%* |
| ***Interviewer reactions to disruption*** | | | | | | | |
| Number of turns with interviewer probing behavior | 10.90 | (0.58) | 11.79 | (0.77) | 9.70 | (0.55) | 2.09** |
| Number of turns with interviewer clarifications | 0.65 | (0.07) | 0.74 | (0.09) | 0.53 | (0.08) | 0.21* |
| Number of turns interviewer gives motivational feedback | 0.55 | (0.06) | 0.62 | (0.07) | 0.47 | (0.07) | 0.15# |
| Number of turns interviewer verifies respondent's answer | 2.54 | (0.73) | 2.67 | (0.72) | 2.36 | (0.89) | 0.31 |
| Number of turns with interviewer disfluencies | 16.80 | (1.87) | 18.09 | (2.46) | 15.07 | (1.90) | 3.02 |
| Number of turns interviewer interrupts respondent | 3.17 | (0.46) | 3.74 | (0.63) | 2.41 | (0.48) | 1.33# |
| Interviewer speed of speech (in words per minute) | 154.55 | (3.20) | 160.52 | (2.51) | 146.60 | (4.60) | 13.92** |
| Percent of interviews with comments about line quality | 7.69% | (1.68%) | 2.86% | (1.25%) | 14.13% | (2.91%) | –11.27%** |
| Percent of interviews with unintelligible interviewer audio | 16.78% | (1.90%) | 17.55% | (2.36%) | 15.76% | (2.53%) | 1.79% |
| ***Dual-frame interviewer speaking behaviors*** | | | | | | | |
| Average dual-frame interviewer speaking duration (in minutes) | 8.56 | (0.19) | 8.39 | (0.19) | 8.76 | (0.22) | –0.37# |
| Average dual-frame interviewer words per minute | 153.62 | (3.65) | 159.28 | (3.23) | 147.11 | (4.58) | 12.17** |

# $p < 0.10$ ; * $p < 0.05$ ; ** $p < 0.01$ ; *** $p < 0.001$

number of conversational turns on which an acceptable answer was given by respondents across all survey questions and the total number of turns with unacceptable answers. Next, we calculate the total number of conversational turns on which a respondent produces at least one disfluency (e.g., "uh" or "um"); says "What?" or "What did you say?"; asks the interviewer to repeat the question, the response options, or a definition that was provided; and interrupts the interviewer. The respondent's speed of speech was calculated by dividing the total number of words spoken by a respondent during the interview by the total number of minutes for all respondent-specific turns, resulting in a measure of words per minute (wpm). We also create an indicator for whether the respondent made any remarks about poor audio quality and being unable to hear (= 1) versus no remarks of this type, as well as a similar indicator variable for interviews for which transcriptionists identified call recordings as having unintelligible audio (= 1) versus no unintelligible audio.

Nine independent variables reflect potential interviewer reactions to disruption (**Table 3**). As with the respondent variables, we calculate the total number of conversational turns on which each interviewer behavior occurs over the entire interview. These measures of interviewer reactions to disruption include the number of turns on which the interviewer engages in probing behavior, provides clarification, provides motivational feedback (e.g., "We really appreciate your answers"), verifies a respondent's answer, has any disfluencies, or interrupts the respondent. We then calculate interviewer speed of speech (in wpm) for all interviewer-specific turns, whether the interview had any interviewer remarks about line quality, and whether transcriptionists identified any instances of unintelligible interviewer audio during the interview. We also calculate the percent of total turns for each behavior by dividing the total number of turns on which each behavior occurred by the total number of turns for that case (**Table 4**).

We also analyze interviewer speaking behaviors for 23 interviewers who conducted both landline and cell phone interviews (who we call dual-frame interviewers).[2] We examine the total amount of time (in minutes) each dual-frame interviewer spent on all interviewer-specific

2. We exclude 104 cases conducted by eight interviewers who only performed cell or landline interviews.

**Table 4.** Percent of conversational turns with interviewer and respondent behaviors

| | *Overall behavior / Total case turns (SE)* | *LL behavior / Total case turns (SE)* | *Cell behavior / Total case turns (SE)* | *LL – Cell diff.* |
|---|---|---|---|---|
| *Behaviors indicating satisficing* | | | | |
| Number of turns respondent agrees with interviewer | 0.10% (0.02%) | 0.12% (0.03%) | 0.06% (0.02%) | 0.06% |
| Number of turns respondent gives DK/REF answers | 0.93% (0.11%) | 1.08% (0.15%) | 0.73% (0.15%) | 0.35%# |
| Number of turns respondent comments about duration | 0.04% (0.01%) | 0.03% (0.01%) | 0.04% (0.01%) | -0.01% |
| Number of turns interviewer comments about duration | 0.23% (0.02%) | 0.21% (0.02%) | 0.24% (0.03%) | -0.03% |
| *Respondent indicators of disruption to perception or comprehension* | | | | |
| Number of turns with an acceptable answer | 33.20% (0.76%) | 32.87% (0.85%) | 33.64% (1.06%) | -0.77% |
| Number of turns with an unacceptable answer | 3.48% (0.21%) | 3.63% (0.25%) | 3.27% (0.29%) | 0.36% |
| Number of turns with respondent disfluencies | 8.42% (0.30%) | 7.72% (0.39%) | 9.35% (0.50%) | -1.63%* |
| Number of turns respondent says "What?" | 0.35% (0.04%) | 0.35% (0.05%) | 0.35% (0.05%) | 0.00% |
| Number of turns respondent asks interviewer to repeat question, definition, or response options | 1.26% (0.09%) | 1.28% (0.12%) | 1.23% (0.10%) | 0.05% |
| Number of turns respondent interrupts interviewer | 2.70% (0.10%) | 2.80% (0.13%) | 2.57% (0.18%) | 0.23% |
| *Interviewer reactions to disruption* | | | | |
| Number of turns with interviewer probing behavior | 5.04% (0.26%) | 5.34% (0.33%) | 4.63% (0.27%) | 0.71%# |
| Number of turns with interviewer clarifications | 0.29% (0.03%) | 0.32% (0.03%) | 0.25% (0.04%) | 0.07% |
| Number of turns interviewer gives motivational feedback | 0.28% (0.03%) | 0.30% (0.04%) | 0.24% (0.03%) | 0.06% |
| Number of turns interviewer verifies respondent's answer | 1.17% (0.31%) | 1.22% (0.31%) | 1.10% (0.39%) | 0.12% |
| Number of turns with interviewer disfluencies | 8.06% (0.88%) | 8.67% (1.17%) | 7.25% (0.85%) | 1.42% |
| Number of turns interviewer interrupts respondent | 1.31% (0.15%) | 1.53% (0.21%) | 1.03% (0.15%) | 0.50%* |

# $p < 0.10$ ; * $p < 0.05$

turns, as well as interviewer speed of speech for this subset of interviewers (detailed analyses for each behavior for this set of interviewers are in Appendix B).

### *Duration of Interview Behaviors*

Finally, we examine the duration of individual behaviors during the interview. These behaviors include the total length of time (in minutes) on all conversational turns made by the interviewer and on all conversational turns made by the respondent, which then were parsed into duration spent on specific interviewer and respondent behaviors. Behaviors also include the amount of time that interviewers spent asking questions (7.14 minutes overall), probing (0.67 minutes), verifying answers (0.08 minutes), clarifying questions (0.04 minutes), and providing feedback (0.57 minutes). Similarly, we examine the amount of

time that respondents in each device spent providing answers overall (3.19 minutes) and by the type of answer that they provided, requesting clarification (0.32 minutes), and asking for feedback (0.29 minutes). We also calculate the percent of total time each behavior takes by dividing the time spent on each behavior for each case by the total interview time of that case.

### *Analysis Methods*

All analyses account for clustering of respondents within interviewers (Hox 1994; Olson and Peytchev 2007; Olson and Bilgen 2011) using the complex survey design procedures (svy procedures) in Stata 14.2 and through multilevel linear regression models (xtmixed procedure) with respondents nested within interviewers.

First, a design-adjusted *t*-test was used to examine differences across devices in overall interview length. Then, we use design-adjusted chi-square tests and *t*-tests to examine whether the composition of the respondent pool and the frequency of each respondent and interviewer behavior differs across device (overall and for dual-frame interviewers alone), a necessary condition for behaviors being able to explain the difference in interview length. Because having only 31 interviewers reduces our degrees of freedom substantially, tests with *p*-values of $0.05 < p < 0.10$ are labeled "marginally statistically significant."

To examine whether respondent characteristics and the frequency of behaviors indicating disruption to perception and comprehension explain differences in interview length across devices, we use multilevel linear regression models predicting log(interview length) with the interviewer as a random effect. First, a base model evaluates the proportion of variance due to interviewers versus respondents in the length of the interview (the intraclass correlation coefficient). Model 1 is the simple bivariate analysis for the unadjusted effect of device (landline/cell phone):

$$\log(interviewminutes) = \beta_{0j} + \beta_1 CellPhone_{ij} + \mu_j + \varepsilon_{ij}$$

for respondent $i$ and interviewer $j$, where $\mu_j \sim N(0, \sigma_\mu^2)$ , and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$.

Blocks of independent variables are then added, with all continuous predictors grand-mean-centered. In model 2, we add interviewer and respondent controls, as well as the respondent characteristics expected to vary across devices. Additional models add measures of satisficing behaviors (model 3), respondent indicators of disruption to the perception and comprehension process (model 4), and indicators of interviewer reactions to disruption (model 5). Thus, the final model is:

$$\begin{aligned}\log(interviewminutes)_{ij} = {} & \beta_0 + \beta_1 CellPhone_{ij} + \sum_{p=1}^{P} \beta_p IwerChar_{pj} \\ & + \sum_{q=1}^{Q} \beta_q RChar_{qij} + \sum_{s=1}^{S} \beta_s SatisficingBeh_{sij} \\ & + \sum_{r=1}^{R} \beta_r RDisruption_{rij} + \sum_{t=1}^{T} \beta_t IDisruption_{tij} + \mu_j + \varepsilon_{ij}\end{aligned}$$

Cohen's $f^2$ effect sizes (Cohen 1988) were calculated for each block of predictors using the method described by Selya et al. (2012). At each step, we examine whether the coefficient for device is reduced in magnitude or significance.

Finally, to examine duration of individual interview behaviors, we examine the average number of minutes that interviewers and respondents spend on different types of interview behaviors. We test whether the length of time spent on these behaviors differs across devices using a design-adjusted *t*-test. For interview behaviors that significantly differ in length across devices, we compare the number of turns on which that behavior occurs, the number of words spoken on these turns, and the actor's speed of speech (in words per second) on these turns across devices.

## Results

### *Differences in Interview Length Between Landline and Cell Phone Respondents*

Cell phone interviews (12.97 minutes) are longer than landline interviews (11.80 minutes) by about 1.17 minutes ($t(30) = 3.40$, $p = 0.002$; **Table 2**). This difference is observed even though we limit length to the time spent asking and answering the same survey questions.

### *Differences in Composition Between Landline and Cell Phone Respondents*

Also shown in **Table 2**, landline respondents tend to be older (60.88 years) than cell respondents (46.88 years, $p < 0.05$) and were less likely to be male (47.35 percent) than cell respondents (70.65 percent, $p < 0.001$). Education levels did not differ across landline and cell respondents.

### *Differences an Interview Behaviors Between Landline and Cell Phone Respondents*

Consistent with previous research, indicators of satisficing from respondent and interviewer behaviors differ only modestly between cell phone and landline respondents (**Table 3**). Differences across devices in agreeing with interviewers and giving don't know/refusal answers (DK/REF) were marginally statistically significant, with respondents on landline phones agreeing more often ($t(30) = -1.78$, $p = 0.085$) and giving more DK/REF answers ($t(30) = -1.94$, $p = 0.061$). There were no differences across the devices in comments about the length of the interview itself.

For respondent indicators of disruption to perception or comprehension, we find that landline and cell phone interviews did not differ in the number of turns on which acceptable or unacceptable answers were provided or the number of turns on which respondents said "What?" or asked for a question, definition, or response options to be repeated. Respondents spoke at a significantly slower rate overall during cell phone interviews (71.46 words per minute, or about 1.2 words per second) than on landline interviews (90.73 words per minute, or about 1.5 words per second) ($t(30) = -8.15$, $p < 0.001$), and had more unintelligible audio (46 percent landline vs. 57 percent cell; $t(30) = 2.27$, $p = 0.030$), although there was no difference in comments about line quality or interruptions. The difference in number of turns with respondent disfluencies was marginally statistically significant, with cell phone interviews having more disfluencies than landline interviews ($t(30) = 1.86$, $p = 0.072$).

Six interviewer indicators of disruption significantly or marginally significantly differed between landline and cell phone interviews.

During landline surveys, interviewers probed more often ($t(30) = -2.93$, $p = 0.006$), provided more clarifications ($t(30) = -2.11$, $p = 0.043$), gave more motivational feedback ($t(30) = -1.85$, $p = 0.075$), and interrupted respondents more often ($t(30) = -1.97$, $p = 0.058$), but there were no differences in verification of respondent answers or in the number of turns with disfluencies across the two devices. Interviewers were more likely to make comments about poor line quality on cell phones (14%) than on landlines (3%; $t(30) = 3.42$, $p = 0.002$), although there is no difference in perceived unintelligible audio from the interviewer side.

Similar to respondents, interviewers spoke at a significantly slower rate during cell phone interviews (146.60 words per minute, or about 2.4 words per second) than the landline interviews (160.52 words per minute, or about 2.7 words per second) ($t(30) = -3.29$, $p = 0.003$). **Table 4** shows that the percent of conversational turns on which a behavior occurs generally does not differ across devices. When there are significant differences, they are less than two percentage points.

To determine whether these differences across devices occur because different interviewers are conducting the interviews or because the same interviewers are behaving differently across devices, we restrict our analysis to only include dual-frame interviewers (bottom panel, **Table 3**). The difference in interviewer speaking duration across devices is marginally significant—cell phone calls last 8.76 minutes versus 8.39 minutes for landline calls ($t(22) = 2.06$, $p = 0.051$). Additionally, dual-frame interviewers speak notably slower during cell phone calls (147.11 wpm) than during landline calls (159.28 wpm, $t(22) = -3.20$, $p < 0.01$). Thus, the same interviewers have slower speech and spend slightly more time administering the survey for cell phone compared to landline respondents.

### *Modeling Interview Length with Respondent Characteristics and Frequencies of Interview Behaviors*

**Table 5** presents coefficients and standard errors for the multilevel models examining whether the frequency of these interviewer and respondent characteristics and behaviors explains the differences across devices in interview length. In the base model, 15.3 percent of the variance in interview duration is due to interviewers (interviewer

**Table 5.** Hierarchical linear regression coefficients and standard errors predicting log(interview length) by device, respondent characteristics, interviewer characteristics, and interview behaviors

| | *Model 1* | | *Model 2* | | *Model 3* | | *Model 4* | | *Model 5* | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* |
| *Device type* | | | | | | | | | | |
| Landline (reference) | – | | – | | – | | – | | – | |
| Cell phone | 0.077*** | *(0.021)* | 0.114*** | *(0.022)* | 0.127*** | *(0.021)* | 0.102*** | *(0.016)* | 0.070*** | *(0.013)* |
| *Respondent characteristics* | | | | | | | | | | |
| Age | – 0.003*** | *(0.001)* | 0.003*** | *(0.001)* | 0.002*** | *(0.000)* | 0.001*** | *(0.000)* | | |
| Education | | | | | | | | | | |
| High school graduate or less (reference) | – | | – | | – | | – | | – | |
| Some college | – | | -0.010 | *(0.026)* | 0.008 | *(0.024)* | -0.012 | *(0.017)* | 0.006 | *(0.014)* |
| College graduate + | – | | -0.050* | *(0.022)* | -0.030 | *(0.021)* | -0.022 | *(0.015)* | -0.007 | *(0.012)* |
| Sex | | | | | | | | | | |
| Male | – | | – | | – | | – | | – | |
| Female | – | | 0.022 | *(0.019)* | 0.005 | *(0.017)* | 0.016 | *(0.012)* | 0.010 | *(0.010)* |
| *Satisficing behaviors* | | | | | | | | | | |
| # turns R agrees with interviewer | – | | – | | 0.039** | *(0.014)* | 0.021* | *(0.010)* | 0.014# | *(0.008)* |
| # turns R gives DK/REF answers | – | | – | | 0.013*** | *(0.002)* | 0.011*** | *(0.002)* | 0.003* | *(0.002)* |
| # turns R comments about duration | – | | – | | -0.027 | *(0.022)* | -0.026# | *(0.015)* | -0.011 | *(0.012)* |
| # turns Iwer comments about duration | – | | – | | 0.023# | *(0.013)* | 0.011 | *(0.009)* | 0.014# | *(0.008)* |
| *Difficulty with perception/comprehension* | | | | | | | | | | |
| Respondent behaviors | | | | | | | | | | |
| # turns – acceptable answer | – | | – | | – | | 0.004*** | *(0.001)* | -0.002 | *(0.001)* |
| # turns – unacceptable answer | – | | – | | – | | 0.007*** | *(0.001)* | 0.000 | *(0.001)* |
| # turns - disfluencies | – | | – | | – | | 0.005*** | *(0.001)* | 0.003*** | *(0.000)* |
| # turns R says "What?" | – | | – | | – | | 0.007 | *(0.005)* | -0.005 | *(0.004)* |
| # turns R asks Iwer to repeat | – | | – | | – | | 0.011*** | *(0.002)* | 0.002 | *(0.002)* |
| # turns R interrupts Iwer | – | | – | | – | | 0.006*** | *(0.001)* | 0.003* | *(0.001)* |
| Speed of speech (in words per minute) | – | | – | | – | | -0.001* | *(0.000)* | 0.000 | *(0.000)* |
| Any comments - line quality during interview | – | | – | | – | | -0.024 | *(0.024)* | -0.032 | *(0.020)* |
| Any unintelligible R audio | – | | – | | – | | 0.046*** | *(0.013)* | 0.018# | *(0.010)* |

| | *Model 1* | | *Model 2* | | *Model 3* | | *Model 4* | | *Model 5* | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* | *coef* | *(SE)* |
| Interviewer behaviors | | | | | | | | | | |
| # turns – probing behavior | – | | – | | – | | – | | 0.010*** | *(0.001)* |
| # turns - clarifications | – | | – | | – | | – | | 0.020*** | *(0.004)* |
| # turns Iwer gives motivational feedback | – | | – | | – | | – | | 0.007 | *(0.008)* |
| # turns Iwer verifies respondent's answer | – | | – | | – | | – | | -0.001 | *(0.002)* |
| # turns - disfluencies | – | | – | | – | | – | | 0.002*** | *(0.001)* |
| # turns interviewer interrupts R | – | | – | | – | | – | | 0.001 | *(0.001)* |
| Speed of speech (in words per minute) | – | | – | | – | | – | | -0.005*** | *(0.000)* |
| Any comments - line quality during interview | – | | – | | – | | – | | 0.024 | *(0.019)* |
| Any unintelligible Iwer audio | – | | – | | – | | – | | -0.004 | *(0.013)* |
| *Interviewer Controls* | | | | | | | | | | |
| Age | – | | 0.003** | *(0.001)* | 0.002* | *(0.001)* | 0.002* | *(0.001)* | 0.000 | *(0.001)* |
| Sex | | | | | | | | | | |
| Male | – | | – | | – | | – | | – | |
| Female | – | | -0.074* | *(0.029)* | -0.061* | *(0.027)* | -0.059* | *(0.028)* | -0.030# | *(0.017)* |
| Tenure | – | | -0.018** | *(0.007)* | -0.015* | *(0.007)* | -0.011 | *(0.007)* | 0.006 | *(0.004)* |
| *Respondent Controls* | | | | | | | | | | |
| Annual income | | | | | | | | | | |
| <$55,000 | – | | – | | – | | – | | – | |
| $55,000 or more | – | | 0.008 | *(0.020)* | -0.002 | *(0.018)* | -0.006 | *(0.013)* | 0.011 | *(0.011)* |
| Race | | | | | | | | | | |
| White, Non-Hispanic | – | | – | | – | | – | | – | |
| All other categories | – | | 0.027 | *(0.024)* | 0.035 | *(0.023)* | 0.033* | *(0.016)* | 0.024# | *(0.013)* |
| Intercept | 2.455*** | *(0.018)* | 2.494*** | *(0.031)* | 2.477*** | *(0.029)* | 2.463*** | *(0.026)* | 2.457*** | *(0.018)* |
| *n* | 429 | | 429 | | 429 | | 429 | | 429 | |
| AIC | -189.304 | | -216.412 | | -273.087 | | -539.299 | | -732.932 | |
| ICC | 0.125 | | 0.070 | | 0.076 | | 0.212 | | 0.081 | |
| Interviewer variance | 0.005 | *(0.002)* | 0.002 | *(0.001)* | 0.002 | *(0.001)* | 0.004 | *(0.001)* | 0.001 | *(0.000)* |
| Residual variance | 0.034 | *(0.002)* | 0.032 | *(0.002)* | 0.027 | *(0.002)* | 0.013 | *(0.001)* | 0.009 | *(0.001)* |

# $p < 0.10$ ; * $p < 0.05$ ; ** $p < 0.01$ ; *** $p < 0.001$

variance = 0.0063; residual variance = 0.0348). As expected, device type is a significant predictor of interview length (model 1 coef. = 0.077, $p < 0.001$) and explains about 3 percent of the respondent-level variance in interview length and about 17 percent of the (very small) interviewer-level variance.

Adding interviewer characteristics (model with only interviewer and respondent controls not shown) reduces the interviewer-level variance by about half. Interviewers who are older (coef. = 0.003, $p = 0.002$), male (coef. = −0.066, $p = 0.032$), and newer to the job (coef. = −0.017, $p = 0.024$) take longer to administer the survey. Neither respondent income nor race are significant predictors of interview duration.

Adding respondent age, education, and sex to the model explains 8.57 percent of the original residual variance in interview length ($f^2 = 0.079$, model 2). As expected, older respondents take longer to complete the interview (coef. = 0.003, $p < 0.001$). Respondents with at least a college degree had shorter interviews (coef. = −0.050, $p = 0.025$) than respondents with a high school degree or less, although this disappears once we account for behaviors during the interview. There is not a significant association between interview length and respondent sex. Including respondent and interviewer characteristics in the model increases the device type coefficient (coef. = 0.114, $p < 0.001$) by 148 percent from its model 1 value, indicating that differences in respondent and interviewer demographics across devices mask rather than account for the disparity in interview length across devices.

Adding satisficing behaviors (model 3) to the model explains an additional 14.29 percent of the residual variance in interview length ($f^2 = 0.166$). Agreeing with the interviewer more often (coef. = 0.039, $p = 0.004$) and providing more DK/REF answers (coef. = 0.013, $p < 0.001$) are statistically significant predictors of interview duration (i.e., longer duration), and more interviewer comments about duration is a marginally statistically significant predictor (coef. = 0.023, $p = 0.076$). When adding satisficing behaviors as predictors, device type continues to predict interview length, with the coefficient again growing stronger (coef. = 0.127, $p < 0.001$; 165 percent of its model 1 value). Thus, differences in satisficing behaviors across the two

devices do not explain why cell phone interviews take longer than landline interviews.

Respondent behaviors indicating difficulty with perception and comprehension account for an additional 40.00 percent of the variance in interview duration ($f^2 = 1.052$, model 4). Seven of the nine behaviors are significant predictors of interview length: number of turns with an acceptable answer (coef. = 0.004, $p < 0.001$), number of turns with an unacceptable answer (coef. = 0.007, $p < 0.001$), number of turns with respondent disfluencies (coef. = 0.005, $p < 0.001$), number of turns on which the respondent asks the interviewer to repeat the question (coef. = 0.011, $p < 0.001$), number of turns on which the respondent interrupts the interviewer (coef. = 0.006, $p < 0.001$), respondent speed of speech (coef. = −0.001, $p = 0.044$), and any unintelligible respondent utterances during an interview (coef. = 0.046, $p < 0.001$). Here, the device type coefficient is still significant (coef. = 0.102, $p < 0.001$; 133 percent of its model 1 value), indicating that differences in the frequency of respondent behaviors related to perception and comprehension do not account for the difference in interview duration across devices.

Model 5 adds interviewer reactions to disruption to the model (11.43 percent of additional variance explained; $f^2 = 0.553$). Interviews with more interviewer turns containing probing behavior (coef. = 0.010, $p < 0.001$), question clarification (coef. = 0.020, $p < 0.001$), and disfluencies (coef. = 0.002, $p < 0.001$) are longer. Interviews during which interviewers speak more quickly are shorter (coef. = −0.005, $p < 0.001$). The device type coefficient is still significant, but reduces to 91 percent of its model 1 value (coef. = 0.070, $p < 0.001$). Thus, differences in the frequency of interviewer behaviors reflecting reaction to perception and comprehension difficulties explain some of the differences across devices, but do not fully account for difference in interview duration.

In sum, our theoretically guided proxy variables of differences in composition of the respondent pool and frequency of behaviors that indicate disruption are predictors of interview length, but largely fail to explain the difference in interview length across landline and cell phones. We now turn to an alternative analysis: the duration of these behaviors.

### *Duration of Interview Behaviors*

We start by decomposing interview length into two parts: the total number of minutes spent during an interview on conversational turns by the interviewer and by the respondent. As shown in **Table 6**, on average, interviewers talk for a total of 8.74 minutes during cell phone interviews, 0.43 minutes longer than landline interviews (8.31 minutes; $t(30) = 2.14$, $p = 0.040$). Thus, about 37 percent of the difference in length between landline and cell phone interviews occurs because of interviewer-related actions. We further partition the length of interviewer turns by the type of behavior performed. There are no significant differences across devices in the amount of time interviewers spend probing, verifying answers, providing clarification, or giving feedback. However, interviewers spend longer (0.37 minutes) asking questions during cell phone interviews (7.35 minutes) than during landline interviews, a marginally statistically significant difference (6.98 minutes; $t(30) = 1.93$, $p = 0.063$), accounting for approximately 32 percent of the difference in length between devices.

Next, we dig deeper into these times to explore why there are differences in the duration of behaviors across devices. Surprisingly, these differences do not arise because interviewers are spending more conversational turns asking questions or saying more words during the cell phone interviews. **Table 7** shows that interviewers use about one more conversational turn to ask questions when a respondent is using a landline phone ($t(30) = -2.96$, $p = 0.006$), and speak approximately 28 more words total (a marginally significant difference) when asking questions during landline interviews ($t(30) = -1.89$, $p = 0.068$). This is in the opposite direction we would expect if it was the number of behaviors themselves driving the difference in interview length. This means that interviewers spend more time asking questions on cell phone interviews because they are speaking more slowly while asking questions, not because they are talking more. In particular, interviewers on landline telephones ask questions at a rate of 2.64 words per second compared to a rate of 2.47 words per second on cell phones, a statistically significant difference ($t(30) = -2.36$, $p = 0.025$).

Similarly, respondents speak for 4.23 minutes during a cell phone interview, 0.74 minutes longer than during a landline interview (3.49 minutes; $t(30) = 3.87$, $p = 0.001$; **Table 6**). The length of time

**Table 6.** Average number of minutes for different types of interviewer and respondent behaviors by device

| *Interview length component (in minutes)* | *Overall mean (SE)* | *Landline mean (SE)* | *Cell mean (SE)* | *Mean LL – Mean cell diff.* | *Overall % of total case duration (SE)* | *LL % of total case duration (SE)* | *Cell % of total case duration (SE)* | *% LL – % Cell diff.* |
|---|---|---|---|---|---|---|---|---|
| Total interview length | 12.30 (0.23) | 11.80 (0.21) | 12.97 (0.32) | –1.17** | – | – | – | – |
| Total length of interviewer turns | 8.49 (0.16) | 8.31 (0.15) | 8.74 (0.22) | –0.43* | 70.15% (0.49%) | 71.55% (0.54%) | 68.27% (0.68%) | 3.28%*** |
| *Length of interviewer turns by behavior* | | | | | | | | |
| Question asking | 7.14 (0.14) | 6.98 (0.12) | 7.35 (0.21) | –0.37# | 59.68% (0.61%) | 60.97% (0.71%) | 57.96% (0.83%) | 3.01%** |
| Probing | 0.67 (0.03) | 0.69 (0.05) | 0.64 (0.04) | 0.05 | 5.06% (0.22%) | 5.39% (0.33%) | 4.63% (0.23%) | 0.76%# |
| Verification | 0.08 (0.02) | 0.07 (0.02) | 0.08 (0.03) | –0.01 | 0.59% (0.15%) | 0.58% (0.14%) | 0.59% (0.19%) | –0.01% |
| Clarification | 0.04 (0.01) | 0.04 (0.01) | 0.04 (0.01) | 0.00 | 0.28% (0.04%) | 0.31% (0.05%) | 0.25% (0.04%) | 0.06% |
| Feedback | 0.57 (0.06) | 0.52 (0.05) | 0.63 (0.09) | –0.11 | 4.53% (0.43%) | 4.30% (0.43%) | 4.84% (0.66%) | –0.54% |
| Total length of respondent turns | 3.81 (0.11) | 3.49 (0.10) | 4.23 (0.16) | –0.74** | 29.84% (0.49%) | 28.40% (0.53%) | 31.76% (0.67%) | –3.36%*** |
| *Length of respondent turns by behavior* | | | | | | | | |
| Answer provided | 3.19 (0.08) | 2.91 (0.07) | 3.57 (0.12) | –0.66*** | 25.40% (0.38%) | 24.07% (0.38%) | 27.17% (0.52%) | –3.10%*** |
| Adequate answer | 2.39 (0.06) | 2.12 (0.04) | 2.76 (0.08) | –0.64*** | 19.53% (0.38%) | 18.06% (0.38%) | 21.48% (0.48%) | –3.42%*** |
| Elaborates answer | 0.27 (0.02) | 0.26 (0.03) | 0.28 (0.04) | –0.02 | 1.87% (0.16%) | 1.90% (0.19%) | 1.83% (0.22%) | 0.07% |
| No elaboration | 2.13 (0.05) | 1.86 (0.03) | 2.49 (0.06) | –0.63*** | 17.66% (0.36%) | 16.16% (0.35%) | 19.65% (0.48%) | –3.49%*** |
| Qualified answer | 0.19 (0.02) | 0.18 (0.02) | 0.20 (0.03) | –0.02 | 1.43% (0.11%) | 1.41% (0.13%) | 1.46% (0.19%) | –0.05% |
| Uncodable answer | 0.46 (0.03) | 0.46 (0.04) | 0.46 (0.05) | 0.00 | 3.36% (0.18%) | 3.42% (0.22%) | 3.29% (0.31%) | 0.13% |
| Don't know | 0.12 (0.02) | 0.13 (0.02) | 0.11 (0.03) | 0.02 | 0.86% (0.10%) | 0.99% (0.13%) | 0.69% (0.15%) | 0.30% |
| Refuse to answer | 0.02 (0.01) | 0.02 (0.00) | 0.03 (0.01) | –0.01# | 0.16% (0.04%) | 0.13% (0.03%) | 0.20% (0.05%) | –0.07%# |
| Clarification | 0.32 (0.02) | 0.31 (0.03) | 0.34 (0.03) | –0.03 | 2.43% (0.15%) | 2.42% (0.20%) | 2.44% (0.20%) | –0.02% |
| Feedback | 0.29 (0.03) | 0.27 (0.04) | 0.31 (0.05) | –0.04 | 1.95% (0.17%) | 1.89% (0.21%) | 2.02% (0.26%) | –0.13% |

#$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

**Table 7.** Mean interviewer question asking behaviors, and respondent answering behaviors by device

| *Variable* | *Overall mean (SE)* | *Landline mean (SE)* | *Cell mean (SE)* | *LL – Cell diff.* |
|---|---|---|---|---|
| *Interviewer – Question asking turns* | | | | |
| Number of turns with question asking | 70.85 (0.17) | 71.20 (0.24) | 70.37 (0.19) | 0.83** |
| Number of words spoken on all question asking turns | 1086.31 (9.66) | 1098.51 (10.99) | 1070.05 (13.20) | 28.46# |
| Average words per second on question asking turns | 2.57 (0.05) | 2.64 (0.04) | 2.47 (0.08) | 0.17* |
| *Respondent – Adequate answer without elaboration turns* | | | | |
| Number of turns without elaboration | 62.57 (0.76) | 62.46 (0.82) | 62.73 (1.14) | −0.27 |
| Number of words spoken on all turns without elaboration | 138.18 (1.80) | 137.87 (2.09) | 138.59 (2.47) | −0.72 |
| Average words per second on turns without elaboration | 1.15 (0.03) | 1.30 (0.02) | .96 (0.02) | 0.34*** |

# $p < 0.10$ ; * $p < 0.05$ ; ** $p < 0.01$ ; *** $p < 0.001$

respondents asked for clarification or gave feedback to the interviewer did not differ across devices. However, respondents spent 0.66 minutes longer answering questions during cell phone interviews (3.57 minutes) than landline interviews (2.91 minutes; $t(30) = 4.68$, $p < 0.001$). Looking at the type of answer provided in more detail, we find that the amount of time spent on qualified answers, uncodable answers, and DK/REF answers did not differ across devices; refusals lasted less than a second longer during cell phone interviews (a marginally significant difference). However, adequate answers on cell phone interviews (2.76 minutes) lasted 0.64 minutes longer than on landlines (2.12 minutes), accounting for 55 percent of the 1.17-minute interview difference between devices ($t(30) = 8.78$, $p < 0.001$). This difference in length is driven by turns on which respondents gave an adequate answer without providing detailed elaboration on their answer; the length of time that respondents spent elaborating their answer did not differ across devices.

As summarized in **Table 7**, the number of turns on which respondents provided an adequate answer without elaboration does not differ across devices, nor does the total number of words spoken on these turns. Just as with interviewers, the speed of speech when answering questions without elaborating on those answers is slower for cell phone interviews (0.96 wps) than for landline interviews (1.30 wps, $t(30) = -11.31$, $p < 0.001$). That is, respondents spend more time answering questions on cell phones not because they are saying more, but because they are talking slower.

## Discussion

In this study, we used behavior codes to explore why cell phone interviews last longer than landline interviews. This is the first study of which we are aware that compares the frequency and duration of interviewer and respondent behaviors across cell phone and landline telephone interviews to try to explain this difference. There are four takeaway messages from this analysis.

First, a difference in interview length persists even after eliminating time related to non-survey activities (e.g., allowing respondents time to find a quiet place to respond) and skip patterns.

Second, respondent characteristics that differ between devices do not drive the disparity in interview length. Although older respondents in this study have longer interviews (consistent with previous research), our cell phone sample was younger than the landline sample. Although it is possible that some unmeasured respondent characteristics may contribute to cell phone interviews lasting longer (e.g., a better measure of "optimizing"), the data from our study demonstrate that cell phone interviews last longer despite commonly hypothesized demographic compositional differences across devices, not because of them.

Third, the number of conversational turns containing respondent indicators of satisficing do not explain the difference in length. Giving more DK/REF responses does lead to longer interviews, and may encourage interviewers to probe more often, thereby increasing interview length, but these behaviors occurred more often during landline interviews, not cell phone interviews.

Fourth, the difference in interview length is best explained by longer interview behaviors during cell phone surveys. Specifically, we find that cell respondents take longer to answer questions, even though the total number and proportion of conversational turns with answers is identical across cell and landline surveys. We also find that the difference in interview length between devices is partially accounted for by interviewers taking longer to ask questions during cell phone surveys, despite having more question-asking turns during landline surveys.

Other research has shown that talkers alter their conversation style if they believe their speaking partner is having trouble understanding

(Giles, Coupland, and Coupland 1991; Clark 1996). It is possible that interviewers notice behaviors suggesting that cell respondents are having difficulty with the survey questions and slow their question-asking speech to be accommodating. In fact, indicators of disruption to perception and comprehension, such as problems with line quality (i.e., inaudible utterances and hearing difficulties) and respondent disfluencies,[3] were more prevalent during cell phone interviews, which may have cued interviewers to slow down. However, we do not have a study design that allows us to directly evaluate this causal argument. Such a design would require (1) information starting from first contact (i.e., recruitment) because the accommodation may start very early and (2) random assignment of questions to different positions in the interview to allow for disentangling changes in speech rate due to accommodation from those due to question characteristics like topic, type, and complexity. Future research of this type is needed to better evaluate why interviewers and respondents speak more slowly in cell interviews.

Our study is limited in its observational nature. Respondents were not randomly assigned to interview device, and therefore device effects may be confounded with unmeasured variables in the sample. In future research, respondents could be randomly assigned to devices as done by Kennedy and Everett (2011), although this kind of design is quickly becoming less feasible, as only 39.6 percent of US adults live in households with both devices (Blumberg and Luke 2017).

In addition, we did not collect information on the respondent's physical location during the interview (e.g., in public versus at home), nor could we examine the amount of time spent outside the question-answer portion of the interview here. These contextual variables could provide insights into how differences in speech rate arise between devices. Although we could not evaluate this directly here, whether speech rate is associated with data-quality measures, and whether this relationship differs across cell and landline surveys, is of interest. Finally, future studies could apply voice-quality assessments made by computer programs to interview recordings (Malfait, Berger, and Kastner 2006), providing a more direct metric of line quality.

3. Disfluencies also could indicate a lack of confidence, searching for a word, or planning to say something difficult (Schober et al. 2012).

The present study answers a long-standing question: Why do cell phone interviews last longer than landline interviews? Cell respondents take longer to provide answers to survey questions (potentially because they are compensating for more perception and comprehension difficulties) than landline respondents. Interviewers on cell surveys also speak more slowly (perhaps to be more accommodating), resulting in a longer interview. Unfortunately, there is little that survey practitioners can do about these behaviors to mitigate the difference in interview length across devices. While telephone interviewer supervisors may be tempted to encourage their staff to maintain a consistent (faster) speed of speech to reduce costs, doing so seems unwise, especially since we do not yet understand why interviewers slow down in cell interviews in the first place or whether their slowing helps respondents provide better answers. In addition, asking respondents to think and speak faster may be detrimental to data quality and/or rapport. Thus, while longer cell phone interviews increase costs for survey organizations, this study suggests that this increase is simply the cost of obtaining quality data via cell phones.

**Supplementary Data** follow the **References**.

## References

Berthold, Andre, and Anthony Jameson. 1999. "Interpreting Symptoms of Cognitive Load in Speech Input." In *Proceedings of the Seventh International Conference on User Modeling,* 235–44. Secaucus: Springer.

Blumberg, Stephen J., and Julian V. Luke. 2017. "Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January–June 2017." Available at https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201712.pdf

Brick, J. Michael, Pat D. Brick, Sarah Dipko, Stanley Presser, Clyde Tucker, and Yangyang Yuan. 2007. "Cell Phone Survey Feasibility in the US: Sampling and Calling Cell Numbers versus Landline Numbers." *Public Opinion Quarterly* 71:23–39.

Ceci, Stephen J. 1991. "How Much Does Schooling Influence General Intelligence and Its Cognitive Components? A Reassessment of the Evidence." *Developmental Psychology* 27:703–22.

Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Couper, Mick P., and Frauke Kreuter. 2013. "Using Paradata to Explore Item Level Response Times in Surveys." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176:271–86.

Dijkstra, Wil. 1999. "A New Method for Studying Verbal Interactions in Survey Interviews." *Journal of Official Statistics* 15:67–85.

Fowler, Floyd J. Jr., and Charles F. Cannell. 1996. "Using Behavioral Coding to Identify Cognitive Problems with Survey Questions." In *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by N. Schwarz and S. Sudman, 15–36. San Francisco: Jossey-Bass.

Fricker, Scott, Mirta Galesic, Roger Tourangeau, and Ting Yan. 2005. "An Experimental Comparison of Web and Telephone Surveys." *Public Opinion Quarterly* 69:370–92.

Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. "Accommodation Theory: Communication, Context, and Consequence." In *Contexts of Accommodation*, edited by H. Giles, J. Coupland, and N. Coupland, 1–68. Cambridge: Cambridge University Press.

Gundersen, Daniel A., Randal S. ZuWallack, James Dayton, Sandra E. Echeverria, and Cristine D. Delnevo. 2014. "Assessing the Feasibility and Sample Quality of a National Random- Digit Dialing Cellular Phone Survey of Young Adults." *American Journal of Epidemiology* 179:39–47.

Hammer, Florian, and Peter Reichl. 2005. "How to Measure Interactivity in Telecommunications." Paper presented at the Annual Congress of the Forum for European ICT and Media Professionals, Vienna, Austria.

Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. "Telephone versus 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.

Hox, Joop J. 1994. "Hierarchical Regression Models for Interviewer and Respondent Effects." *Sociological Methods & Research* 22:300–318.

Jenkins, Cleo R., and Don A. Dillman. 1997. "Towards a Theory of Self-Administered Questionnaire Design." In *Survey Measurement and Process Quality*, edited by L. E. Lyberg, et al., 165–96. New York: John Wiley & Sons.

Keeter, Scott, and Courtney Kennedy. 2006. "The Cell Phone Challenge to Survey Research." *Pew Research Center for the People and the Press*. Available at http://www.people-press.org/2006/05/15/the-cell-phone-challenge-to-survey-research/

Kennedy, Courtney. 2010. *Nonresponse and Measurement Error in Mobile Phone Surveys*. Doctoral dissertation, University of Michigan–Ann Arbor.

Kennedy, Courtney, and Stephen E. Everett. 2011. "Use of Cognitive Shortcuts in Landline and Cell Phone Surveys." *Public Opinion Quarterly* 75:336–48.

Kennedy, Courtney, Kyley McGeeney, and Scott Keeter. 2016. "The Twilight of Landline Interviewing." Available at http://www.pewresearch.org/2016/08/01/the-twilight-of-landline-interviewing

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.

Krosnick, Jon A., Allyson L. Holbrook, Matthew K. Berent, Richard T. Carson, W. Michael Hanemann, Raymond J. Kopp, Robert Cameron Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. 2002. "The Impact of 'No Opinion' Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66:371–403.

Kuusela, Vesa, and Veijo Notkola. 1999. "Survey Quality and Mobile Phones." Paper presented at the International Conference on Survey Nonresponse, Portland, OR, USA. Landis, J. Richard, and Gary G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33:159–74.

Lavrakas, Paul J., Stephen Blumberg, Michael Battaglia, John Boyle, Michael Brick, Trent Buskirk, Charles DiSogra, David Dutwin, Mansour Fahimi, Howard Fienberg, Anna Fleeman, Thomas M. Guterbock, John Hall, Scott Keeter, Courtney Kennedy, Michael Link, Linda Piekarski, Charles D. Shuttles, Charlotte Steeh, Trevor Tompson, Randall ZuWallack. 2010. "New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in the US with Respondents Reached via Cell Phone Numbers." Available at https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/2010AAPORCellPhoneTFReport.pdf

Lavrakas, Paul J., Charles D. Shuttles, Charlotte Steeh, and Howard Fienberg. 2007. "The State of Surveying Cell Phone Numbers in the United States 2007 and Beyond." *Public Opinion Quarterly* 71:840–54.

Lavrakas, Paul J., Trevor N. Tompson, Robert Benford, and Christopher Fleury. 2010. "Investigating Data Quality in Cell Phone Surveying." Paper presented at the Annual Conference of the American Association of Public Opinion Research, Chicago, IL, USA.

Linville, Sue E. 2001. *Vocal Aging*. San Diego, CA: Singular Thomson Learning.

Lynn, Peter, and Olena Kaminska. 2012. "The Impact of Mobile Phones on Survey Measurement Error." *Public Opinion Quarterly* 77:586–605.

Malfait, Ludovic, Jens Berger, and Martin Kastner. 2006. "The ITU-T Standard for Single-Ended Speech Quality Assessment." *IEEE Transactions on Audio, Speech, and Language Processing* 14:1924–34.

McGeeney, Kyley. 2016. "Pew Research Center Will Call 75% Cellphones for Surveys in 2016." Available at http://www.pewresearch.org/fact-tank/2016/01/05/pew-research-center-will-call-75-cellphones-for-surveys-in-2016

McHugh, Mary L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22:276–82.

Miller, Kristen, Daniel Mont, Aaron Maitland, Barbara Altman, and Jennifer Madans. 2011. "Results of a Cross-National Structured Cognitive Interviewing Protocol to Test Measures of Disability." *Quality & Quantity* 45:801–15.

Nathan, Gad. 2001. "Telesurvey Methodologies for Household Surveys—A Review and Some Thoughts for the Future." *Survey Methodology* 27:7–32.

Olson, Kristen, and Ipek Bilgen. 2011. "The Role of Interviewer Experience on Acquiescence." *Public Opinion Quarterly* 75:99–114.

Olson, Kristen, and Andy Peytchev. 2007. "Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes." *Public Opinion Quarterly* 71:273–86.

Olson, Kristen, and Jolene D. Smyth. 2015. "The Effect of CATI Questions, Respondents, and Interviewers on Response Time." *Journal of Survey Statistics and Methodology* 3:361–96.

Ongena, Yfke P., and Wil Dijkstra. 2006. "Methods of Behavior Coding of Survey Interviews." *Journal of Official Statistics* 22:419–51.

Peytchev, Andy, and Benjamin Neely. 2013. "RDD Telephone Surveys Toward a Single-Frame Cell-Phone Design." *Public Opinion Quarterly* 77:283–304.

Salthouse, Timothy A. 1991. "Mediation of Adult Age Differences in Cognition by Reductions in Working Memory and Speed of Processing." *Psychological Science* 2:179–83.

Schaeffer, Nora C., and Douglas W. Maynard. 1996. "From Paradigm to Prototype and Back Again: Interactive Aspects of Cognitive Processing in Standardized Survey Interviews." In *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by N. Schwarz and S. Sudman, 65–88. San Francisco: Jossey-Bass.

Schober, Michael F., and Jonathan E. Bloom. 2004. "Discourse Cues That Respondents Have Misunderstood Survey Questions." *Discourse Processes* 38:287–308.

Schober, Michael F., Frederick G. Conrad, Christopher Antoun, Patrick Ehlen, Stefanie Fail, Andrew L. Hupp, Michael Johnston, Lucas Vickers, H. Yanna Yan, and Chan Zhang. 2015. "Precision and Disclosure in Text and Voice Interviews on Smartphones." *PloS One* 10:e0128337.

Schober, Michael F., Fred G. Conrad, Wil Dijkstra, and Yfke. P. Ongena. 2012. "Disfluencies and Gaze Aversion in Unreliable Responses to Survey Questions." *Journal of Official Statistics* 28:555–82.

Schwarz, Norbert, Fritz Strack, Hans J. Hippler, and George Bishop. 1991. "The Impact of Administration Mode on Response Effects in Survey Measurement." *Applied Cognitive Psychology* 5:193–212.

Selya, Arielle S., Jennifer S. Rose, Lisa C. Dierker, Donald Hedeker, and Robin J. Mermelstein. 2012. "A Practical Guide to Calculating Cohen's F2, a Measure of Local Effect Size, From PROC MIXED." *Frontiers in Psychology* 3:1–6.

Thrasher, James F., Anne C. K. Quah, Gregory Dominick, Ron Borland, Pete Driezen, Rahmat Awang, Maizurah Omar, Warwick Hosking, Buppha Sirirassamee, and Marcelo Boado. 2011. "Using Cognitive Interviewing and Behavioral Coding to Determine Measurement Equivalence Across Linguistic

and Cultural Groups: An Example from the International Tobacco Control Policy Evaluation Project." *Field Methods* 23:439–60.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

U.S. Bureau of Labor Statistics. 2016. "Average Hours Per Day Spent in Selected Activities by Age." Available at https://www.bls.gov/charts/american-time-use/activity-by-age.htm

Verhoeven, Jo, Guy De Pauw, and Hanne Kloots. 2004. "Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands." *Language and Speech* 47:297–308.

Vicente, Paula, Elizabeth Reis, and Maria Santos. 2009. "Using Mobile Phones for Survey Research." *International Journal of Market Research* 51:613–33.

Viera, Anthony J., and Joanne M. Garrett. 2005. "Understanding Interobserver Agreement: The Kappa Statistic." *Family Medicine* 37:360–63.

Witt, Lindsey, Randall ZuWallack, and Frederica Conrey. 2009. "Out and About: An Evaluation of Data Quality in Cell Phone Surveys." Paper presented at the Annual Conference of the American Association of Public Opinion Research, Hollywood, FL, USA.

Yont, Kristine M., Lynne E. Hewitt, and Adele W. Miccio. 2002. "'What Did You Say?' Understanding Conversational Breakdowns in Children with Speech and Language Impairments." *Clinical Linguistics & Phonetics* 16:265–85.

Online Appendix A. NOP Questionnaire.

**Q1** I am going to read some various institutions in American society. Would you say that you "do" or "do not" have a lot of confidence in **(read and rotate Q1A-Q1N)**?

1 Do
2 Do not
3 (DK)
4 (Refused)

| | |
|---|---|
| **Q1A** | The President of the U.S. |
| **Q1B** | The U.S. Congress |
| **Q1C** | Police departments and public prosecutors' offices |
| **Q1D** | The courts |
| **Q1E** | The military |
| **Q1F** | Churches |
| **Q1G** | Federal government agencies |
| **Q1H** | Local governments |
| **Q1I** | Schools |
| **Q1J** | Hospitals |
| **Q1K** | Newspapers |
| **Q1L** | TV stations |
| **Q1M** | Major corporations |
| **Q1N** | Labor unions |

**Q2** How would you rate relations between the United States and Japan at present? Would you say they are **(read 5-1)**?

5 Very good
4 Good
3 Just fair
2 Poor, OR
1 Very Poor
6 (DK)
7 (Refused)

**Q3** How much do you trust Japan? Would you say **(read 4-1)**?

4 Very much
3 Some
2 Not very much, OR
1 Not at all
5 (DK)
6 (Refused)

**Q4** Do you think that the relationship between the United States and Japan will **(read 5-1)**?

5 Get much better
4 Get somewhat better
3 Stay the same
2 Get somewhat worse, OR
1 Get much worse
6 (DK)
7 (Refused)

**Q5** In your opinion, how functional is the U.S. political system these days? Do you think the U.S. political system is highly functional, somewhat functional, not very functional or not functional at all?

4 Highly functional
3 Somewhat functional
2 Not very functional
1 Not functional at all
5 (DK)
6 (Refused)

**Q6** Which countries or regions do you think will become a military threat to the United States? How about **(read and rotate Q6A-Q6M)**?

1 Yes
2 No
3 (DK)
4 (Refused)

| | |
|---|---|
| **Q6A** | Japan |
| **Q6B** | South Korea |
| **Q6C** | China |
| **Q6D** | Association of Southeast Asian Nations (ASEAN) |
| **Q6E** | European Union (EU) |
| **Q6F** | Russia |
| **Q6G** | Taiwan |
| **Q6H** | North Korea |
| **Q6I** | India |
| **Q6J** | Middle East |
| **Q6K** | Central and South Pacific nations, such as Australia and New Zealand |
| **Q6L** | Africa |
| **Q6M** | Latin America |

**Q7** To what extent do you think the U.S.-Japan Security Treaty contributes to the security of the Asia-Pacific region? Would you say **(read 4-1)**?

4 Contributes greatly
3 Contributes somewhat
2 Does not contribute very much, OR
1 Does not contribute at all
5 (DK)
6 (Refused)

**Q8** At present, the U.S. maintains many U.S. military bases in Japan under the U.S.-Japan Security Treaty. Do you think the scale of U.S. military presence in Japan should be increased, maintained, reduced, or eliminated altogether?
***(Interviewer:* Read 4-1*)***

4 Should be increased
3 Should be maintained
2 Should be reduced, OR
1 Should be eliminated altogether
5 (DK)
6 (Refused)

**Q9** How would you rate relations between the United States and China at present? Would you say they are **(read 5-1)**?

5 Very good
4 Good
3 Just fair
2 Poor, OR
1 Very Poor
6 (DK)
7 (Refused)

**Q10** How much do you trust China? Would you say **(read 4-1)**?

4 Very much
3 Some
2 Not very much, OR
1 Not at all
5 (DK)
6 (Refused)

**Q11** Which country, Japan or China, do you feel will be more important to the United States in the future for POLITICAL matters?
***(Interviewer:* Read 2-1*)***

2 Japan, OR
1 China
3 (Both equally)
4 (DK)
5 (Refused)

**Q12** Which country, Japan or China, do you feel will be more important to the United States in the future for ECONOMIC matters?
***(Interviewer:* Read 2-1*)***

2 Japan, OR
1 China
3 (Both equally)
4 (DK)
5 (Refused)

**Q13** I am going to read out some issues regarding China. Would you say you are or are not concerned about **(Read and Rotate Q13A-Q13H)**?

1 Concerned, OR
2 Not concerned
3 (DK)
4 Refused

| | |
|---|---|
| **Q13A** | The rapid expansion of China's economy |
| **Q13B** | The valuation of China's currency |
| **Q13C** | The theft of intellectual property such as counterfeit consumer goods |
| **Q13D** | The political regime |
| **Q13E** | Human rights |
| **Q13F** | Strengthened military power |
| **Q13G** | Territorial disputes with China's neighboring countries |
| **Q13H** | China launching a cyber-attack on the U.S. |

**Q14** Given China's increasing influence in the Asia-Pacific region, do you think the U.S. should increase, maintain or reduce its military presence in the Asia-Pacific region? ***(Interviewer:* Read 1 - 3*)***

1 Should increase
2 Should maintain at its current level, OR
3 Should reduce
4 (DK)
5 (Refused)

**Q15** Would you say Japan's international influence has grown stronger, weaker or has it remained the same in recent years?

3 Stronger
2 Weaker
1 Remained the same
4 (DK)
5 (Refused)

**Q16** And how about the U.S.? Would you say The United States' influence in the international community has grown stronger, weaker or, has it remained the same in recent years?

3 Stronger
2 Weaker
1 Remained the same
4 (DK)
5 (Refused)

**Q17** How much do you trust South Korea? Would you say **(read 4-1)**?

4 Very much
3 Some
2 Not very much, OR
1 Not at all
5 (DK)
6 (Refused)

**Q18** Regarding North Korea, which issues should the U.S. and Japanese governments, working in cooperation, give priority to resolving? How about **(read and rotate Q18A-Q18F)**?

1 Yes
2 No
3 (DK)
4 (Refused)

| | |
|---|---|
| **Q18A** | Getting North Korea to abandon its nuclear weapons program |
| **Q18B** | Getting North Korea to end its missile program and its missile launch |
| **Q18C** | Resolving the cases involving the abduction of Japanese citizens by North Korea |
| **Q18D** | Normalizing diplomatic relations between the U.S. and North Korea |
| **Q18E** | Providing economic aid to North Korea |
| **Q18F** | Changing North Korea's political and economic systems |

**Q19** A number of Asia-Pacific countries, including the U.S. and Japan, are now in talks to finalize the Trans-Pacific Partnership Agreement, or TPP. This agreement will liberalize trade and includes the agricultural sector. What impact, if any, do you think the TPP will have on the United States? Do you think it will have a **(read 5-1)**?

5 Very positive impact,
4 Somewhat positive impact,
3 Neither positive nor negative impact
2 Somewhat negative impact, OR
1 Very negative impact
6 Don't know about the TPP
7 (Refused)

**Q20** Which of the following comes closest to your opinion of the future of nuclear power plants in the United States?
***(Interviewer:* Read 1 - 4*)***

1 We should increase the number of nuclear power plants
2 We should maintain the current number of nuclear power plants
3 We should reduce the number of nuclear power plants, OR
4 We should eliminate all nuclear power plants
5 (Other)
6 (DK)
7 (Refused

**DEMOGRAPHICS BEGIN HERE:**

***(Interviewer:* READ:*)***
The following questions are for demographic purposes only.

**D1** What is your age?
***(Interviewer:*** Open ended **and code actual age*)***

00 (Refused)
99 99+

**List Other:**Y

**D2** Are you currently **(read 06-11, then 01)**?

01 OR, something else (list)
02 (DK)
03 (Refused)
04 HOLD
05 HOLD
06 Self-employed
07 A salaried employee
08 A homemaker
09 A student
10 Unemployed
11 Retired

**List Other:**Y

***Skip:* (If code 01-03 in D2, Skip to D5;**
**If code 07 in D2, Skip to D4;**
**If code 08-11 in D2, Skip to D5;**
**Otherwise, Continue)**

**D3** Please select the category that BEST describes your current job. Is it **(read 1-3)**?

1 Agriculture or forestry
2 Commerce, industry, or service industries, OR
3 Freelance
4 (DK)
5 (Refused)

***Skip:* (All in D3, Skip to D5)**

**D4** Please select the category that BEST describes your current job. Is it **(read 1-3)**?

1 Manager or specialist
2 Administrative or technical position, OR
3 Labor or service-related position
4 (DK)
5 (Refused

**D5** What was your annual household income in 2012, before taxes? Was it **(read 01-08)**?

01 Less than $15,000
02 $15,000 to less than $25,000
03 $25,000 to less than $35,000
04 $35,000 to less than $45,000
05 $45,000 to less than $55,000
06 $55,000 to less than $75,000
07 $75,000 to less than $100,000, OR
08 $100,000 or more
09 (DK)
10 (Refused)

**D6** Do you approve or disapprove of the way Barack Obama is handling his job as president?
***(Interviewer:* Read 1-2*)***

1 Approve, OR
2 Disapprove
3 (Neither approve nor disapprove)
4 (DK)
5 (Refused)

**D7** Do you support the Republican Party, the Democratic Party, some other party, or none of them?

01 Some other party (list)
02 (DK)
03 (Refused)
04 None of them
05 HOLD
06 Republican
07 Democratic

**List Other:**Y

**D8** Are you, yourself, of Hispanic origin or descent, such as Mexican, Puerto Rican, Cuban, or other Spanish background?

1 Yes
2 No
3 (DK)
4 (Refused)

**D9** What is your race? Are you White, African-American, Asian, or some other race?

01 Other (Do NOT list)
02 (DK)
03 (Refused)
04 HOLD
05 HOLD
06 White
07 African-American/Black
08 (Hispanic)
09 Asian

**D10** Including yourself, how many adults, 18 years of age or older, live in this household?
***(Interviewer***: **Open ended and code actual number**)

01 01-
96 96
97 97 or more
98 (DK)
99 (Refused)

**D11** What is the highest level of education you have completed?
***(Interviewer:*** **Open ended and code*)***

1 Less than high school graduate (0-11)
2 High school graduate (12)
3 Some college
4 Trade/Technical/Vocational training
5 College graduate
6 Postgraduate work/Degree
7 (DK)
8 (Refused)

***Skip: (If Landline Respondent, Autocode D12A as 1 and Skip to Note before D12B;***
***Otherwise, Continue)***

**D12A** Do you have a working landline telephone in your home?

1 Yes
2 No
3 (DK)
4 (Refused)

***Skip: (If Mobile Respondent, Autocode D12B as 1 and Skip to Note #2 before D14;***
***Otherwise, Continue)***

**D12B** Do you have a working cell phone that you receive and make calls on?

1 Yes
2 No
3 (DK)
4 (Refused)

***Skip: (If code 2-4 in D12B, Continue;***
***Otherwise, Skip to Note #1 before D14)***

**D13** ***(NOTE: This question, D13, was removed from analysis as it was asked uniquely to Landline respondents)***

Does anyone in your household have a working cell phone?
***(Interviewer: This can include children under 18 in the household)***

1 Yes
2 No
3 (DK)
4 (Refused)

***(Programmer: Note #1:)***
***Skip: (If code 2-4 in D13, Skip to Thank and Validate;***
***Otherwise, Continue)***

***(Programmer: Note #2:)***
***Skip: (If code 1 in D12A AND [code 1 in D12B or D13], Continue;***
***Otherwise, Skip to Thank and Validate)***

**D14** Of all the telephone calls your household receives which best describes your household's phone use **(read 1-3)**?

1 All or almost all calls are received on cell phones
2 Some are received on cell phones and some on regular phones, OR
3 Very few or none are received on cell phones
4 (DK)
5 (Refused)

***Skip: (If code 1 in D12B, Continue; Otherwise, Skip to Thank and Validate)***

**D15** How many different residential phone NUMBERS do you have coming into your household, not including lines dedicated to a fax machine, modem, or used strictly for business purposes? Do not include cellular phones.
***(Interviewer:*** Open ended and code***)***

0 Zero
1 One
2 Two
3 Three
4 Four
5 Five or more
6 (DK)
7 (Refused)

Online Appendix B. Summary of Behaviors Indicating Satisficing and Disruption to Perception or Comprehension for Dual-Frame Interviewers

| | Overall Mean | (SE) | Landline Mean | (SE) | Cell Phone Mean | (SE) | LL – Cell Diff. |
|---|---|---|---|---|---|---|---|
| Total Number of Turns | 204.59 | (4.99) | 207.00 | (5.28) | 201.81 | (6.56) | 5.19 |
| Behaviors Indicating Satisficing | | | | | | | |
| Number of turns respondent agrees with interviewer | 0.22 | (0.07) | 0.31 | (0.11) | 0.11 | (0.05) | 0.20* |
| Number of turns respondent gives DK/REF answers | 2.03 | (0.33) | 2.24 | (0.46) | 1.79 | (0.38) | 0.45 |
| Number of turns respondent comments about duration | 0.07 | (0.02) | 0.06 | (0.03) | 0.08 | (0.02) | -0.02 |
| Number of turns interviewer comments about duration | 0.47 | (0.05) | 0.45 | (0.05) | 0.48 | (0.07) | -0.03 |
| Respondent Indicators of Disruption to Perception or Comprehension | | | | | | | |
| Number of turns with an acceptable answer | 65.32 | (0.82) | 66.07 | (0.79) | 64.46 | (1.17) | 1.61 |
| Number of turns with an unacceptable answer | 7.66 | (0.53) | 7.88 | (0.72) | 7.40 | (0.81) | 0.48 |
| Number of turns with respondent disfluencies | 17.55 | (0.63) | 15.98 | (0.95) | 19.36 | (1.09) | -3.38† |
| Number of turns respondent says "What?" | 0.80 | (0.10) | 0.80 | (0.14) | 0.80 | (0.13) | 0.00 |
| Number of turns respondent asks interviewer to repeat question, definition, or response options | 6.39 | (0.56) | 6.78 | (0.79) | 5.94 | (0.54) | 0.84 |
| Number of turns respondent interrupts interviewer | 6.22 | (0.39) | 6.63 | (0.53) | 5.74 | (0.48) | 0.89 |
| Respondent speed of speech (in words per minute) | 81.25 | (1.62) | 89.13 | (2.07) | 72.17 | (1.73) | 16.96*** |
| Percent of interviews with comments about line quality | 7.08% | (1.63%) | 6.90% | (1.59%) | 7.29% | (2.15%) | -0.39% |
| Percent of interviews with unintelligible respondent audio | 49.23% | (3.61%) | 41.95% | (4.08%) | 57.62% | (3.98%) | -15.67%** |
| Interviewer Reactions to Disruption | | | | | | | |
| Number of turns with interviewer probing behavior | 11.21 | (0.71) | 12.34 | (0.95) | 9.90 | (0.66) | 2.44** |
| Number of turns with interviewer clarifications | 0.69 | (0.09) | 0.79 | (0.12) | 0.58 | (0.09) | 0.21† |
| Number of turns interviewer gives motivational feedback | 0.57 | (0.07) | 0.66 | (0.08) | 0.47 | (0.07) | 0.19* |
| Number of turns interviewer verifies respondent's answer | 2.65 | (0.95) | 2.94 | (0.95) | 2.32 | (1.07) | 0.62 |
| Number of turns with interviewer disfluencies | 18.51 | (2.26) | 21.47 | (2.74) | 15.11 | (2.10) | 6.36** |
| Number of turns interviewer interrupts respondent | 3.42 | (0.58) | 4.08 | (0.86) | 2.66 | (0.56) | 1.42 |
| Interviewer speed of speech (in words per minute) | 153.62 | (3.65) | 159.28 | (3.23) | 147.11 | (4.58) | 12.17** |
| Percent of interviews with comments about line quality | 6.77% | (1.73%) | 2.30% | (1.07%) | 11.92% | (2.96%) | -9.62%** |
| Percent of interviews with unintelligible interviewer audio | 16.00% | (2.24%) | 16.67% | (2.91%) | 15.23% | (2.68%) | 1.44% |

Note: †$p$<.10, *$p$<.05, **$p$<.01, ***$p$<.001