Bias in child maltreatment self-reports using Interactive Voice Response

Nancy J. Kepple,
Luskin School of Public Affairs, Department of Social Welfare at University of California, Los Angeles

Bridget Freisthler, and
Luskin School of Public Affairs, Department of Social Welfare at University of California, Los Angeles

Michelle Johnson-Motoyama
School of Social Welfare at University of Kansas

Abstract

Few methods estimate the prevalence of child maltreatment in the general population due to concerns about socially desirable responding and mandated reporting laws. Innovative methods, such as Interactive Voice Response (IVR), may obtain better estimates that address these concerns. This study examined the utility of Interactive Voice Response (IVR) for child maltreatment behaviors by assessing differences between respondents who completed and did not complete a survey using IVR technology. A mixed-mode telephone survey was conducted in English and Spanish in 50 cities in California during 2009. Caregivers (n = 3,023) self-reported abusive and neglectful parenting behaviors for a focal child under the age of 13 using Computer-Assisted Telephone Interviewing and IVR. We used Hierarchical Generalized Linear Models to compare survey completion by caregivers nested within cities for the full sample and age-specific ranges. For demographic characteristics, caregivers born in the United States were more likely to complete the survey when controlling for covariates. Parenting stress, provision of physical needs, and provision of supervisory needs were not associated with survey completion in the full multivariate model. For caregivers of children 0 to 4 years (n = 838), those reporting they could often or always hear their child from another room had a higher likelihood of survey completion. The findings suggest IVR could prove to be useful for future surveys that aim to estimate abusive and/or neglectful parenting behaviors given the limited bias observed for demographic characteristics and problematic parenting behaviors. Further research should expand upon its utility to advance estimation rates.

Keywords

child maltreatment; survey methods; general population estimates; interactive voice response

Please address correspondence to: Nancy J. Kepple, M.S.W., Department of Social, Welfare at the UCLA Luskin School of Public Affairs, 3250 Public Affairs Building, Box, 951656, Los Angeles, CA 90095-1656; phone: (310) 755-8802; fax: (310) 206-7564; njwilliams@ucla.edu.

Rates of substantiated child maltreatment in the United States have demonstrated modest declines over the past decade yet remain at concerning levels (Child Trends, 2012). In 2011, child protective service systems identified 9.1 per 1,000 children to be victims of abuse or neglect (US DHHS, 2012). However, results from the small number of general population surveys estimate much higher rates of child maltreatment than those captured by child protective service response systems (Finkelhor, Turner, Ormrod, & Hamby, 2009; Hussey, Chang, & Kotch, 2006; Sedlak et al., 2010; Straus et al., 1998) General population estimates obtain more accurate estimates of child maltreatment by overcoming the limitations of administrative data, which depend primarily on surveillance and reporting, contain limited demographic information, and often suffer from agency-level data gaps and errors (Drake & Jonson-Reid, 1999; Wulczyn, 2009). General population surveys have typically used methods such as sentinel reporters (i.e. community professionals who encounter children and families as a part of their job) or victim recall of childhood experiences (Finkelhor et al., 2009; Hussey et al., 2006; Sedlak et al., 2010). While these methods are preferable to administrative data sources, several limitations remain such as sentinel reporters' ability to accurately identify children at risk for maltreatment or unreliable self-reporting of early life events (Hardt & Rutter, 2004; Sedlak & Ellis, 2014). It is less common for general population surveys to acquire caregiver self-report of maltreatment behaviors due to concerns about the potential under-reporting of these behaviors likely due to respondent fear of disapproval from the interviewer and/or being reported to child protective services for truthful responding (Cicchetti & Toth, 2005; Tourangeau & Smith, 1996).

Telephone surveys provide an economical option to directly sample caregivers from the general population across large geographic areas and often produce higher quality data due to lower rates of item non-response when compared to mail or web surveys (Bowling, 2005; Lesser, Newton, & Yang, 2012). Innovative telephone survey methods, such as Interactive Voice Response (IVR), has potential for advancing our understanding of child maltreatment by obtaining general population estimates in a way that addresses potential bias in self-reporting. IVR is a computerized interviewing system that plays a recording of the questions over the phone and relies on touch-tone entry by respondents to record their answers (Tourangeau, Steiger, & Wilson, 2002). This technology differs from the more frequently used Computer-Assisted Telephone Interviewing (CATI) methods, which depends on live interviewers to read prompts and questions from a computer program and enter respondent answers directly into the same program (Bowling, 2005).

Prior studies have demonstrated that IVR minimizes socially desirable responding for topics such as alcohol/drug use and sex-related behaviors (e.g., Midanik & Greenfield, 2008; Schroder, Johnson, & Wiebe, 2007; Turner et al., 1998), resulting in higher rates of disclosure for socially undesirable behaviors when compared to the use of a live interviewer using CATI methods (Midanik & Greenfield, 2008; Tourangeau & Smith, 1996). This benefit of IVR is comparable to benefits observed with corresponding in-person survey strategies (Beach et al., 2010). These observed differences are likely due to respondent's increased perception of confidentiality and lower levels of discomfort in disclosing sensitive information with an automated system (Corkrey & Parkinson, 2002b; Groves, Cialdini, & Couper, 1992; Kreuter, Presser, & Tourangeau, 2008). Therefore, IVR methods may produce better estimates of maltreatment behaviors in the general population by addressing

biases associated with social desirability. However, this benefit must be balanced with the lower survey completion rates observed with the use of IVR when compared to the use of a live interviewer (Rodriguez et al., 2006). The automated IVR system lacks the psychological barriers to dropping out that can be provided by a live interviewer who can motivate and persuade a respondent to complete the survey (Groves et al., 1992). As a result, survey dropout rates with IVR can be substantial, typically ranging from 5% to 45% (Galesic, Tourangeau, & Couper, 2006; Tourangeau et al., 2002). Therefore, the benefits of IVR for eliciting responses to questions on sensitive topics may be negated if survey responses are biased due to differential dropout rates among respondents, particularly among those who engage in behaviors associated with child maltreatment.

Factors Associated with Survey Completion

Multiple factors influence respondent survey completion. In general, respondent behavior can be influenced by one's reaction to the survey modality (e.g., presence of an interviewer), experience of respondent fatigue towards the end of a survey (e.g., being tired or bored with the survey), one's cognitive reaction to survey items (e.g., difficulty comprehending the question and/or response options), and/or one's emotional reaction to survey items (e.g., respondent discomfort), all of which can result in higher dropout rates (e.g., Galesic et al., 2006; Tourangeau et al., 2002). As stated earlier, IVR approaches do not utilize live interviewers who provide barriers to dropping out because of psychological factors such as authority (e.g., people usually find it rude to hang-up on an interviewer once engaged) and reciprocity (e.g., interviewers can provide additional encouragement and/or feedback to respondents to keep them engaged in the process) (Groves et al., 1992). Respondent fatigue for longer IVR surveys may also increase dropout rates due the lack of interviewer barriers and/or respondent boredom with an automated system (Galesic et al., 2006). Survey completion patterns observed with IVR typically result an initial drop-out during transition to the automated system continued by drop-out throughout the survey (compared to only an initial drop-out observed with CATI), suggesting respondent fatigue and/or reaction to the IVR modality may result when interviewer barriers are removed (Galesic et al., 2006; Kreuter et al., 2008; Tourangeau et al., 2002).

Survey completion studies have typically assessed specific respondent characteristics, such as respondent age, gender, and/or income; however, the studies have not observed consistent findings based on demographics alone (Groves & Couper, 1998; Groves et al., 1992). In their review of IVR studies, Corkrey & Parkinson (2002a) also suggest there is little evidence that the use of IVR methods alone result in biased demographic characteristics. Overall, the relative importance of demographic characteristics in survey completion may be associated with the survey topic (Groves et al., 1992; Groves & Couper, 1998). For example, a survey on the usefulness of subsidized student loans may result in a greater dropout rate for populations with limited interest in the topic.

Moreover, respondents may be more likely to dropout when they experience extra burden or demand, including difficulty aurally processing questions or discomfort when answering questions about taboo and/or illegal behaviors (Bowling, 2005; Peytchev, 2009). For example, respondents with difficulty processing information independently (e.g., lower IQ

or language barriers) may have a higher likelihood of dropping out of a survey due to the higher cognitive burden of IVR, which requires participants to aurally process the information without assistance (Bowling, 2005; Peytchev, 2009). Alternatively, caregivers of children who highly identify with their parenting role may be motivated to complete a parenting survey; however, this motivation can be undermined by the survey if it causes respondents to feel uneasy (Crouper & Groves, 1996). The burden of answering sensitive questions often can lead to item nonresponse in surveys due to the respondent's discomfort (Bosnjak & Tuten, 2001). The combination of this type of burden with the lack of barriers to drop-out in an IVR section may result in lower completion rates for caregivers who endorse maltreatment behaviors and subsequently bias survey outcomes.

However, IVR research has yet to examine survey completion behavior on the topic of neglectful and abusive parenting behaviors. It is important to explore how key demographic and parenting behaviors may differ between respondents who choose to complete an IVR survey from those who do not in order to gauge the usefulness of IVR methods for estimating neglectful and abusive behaviors in the general population. Differences in survey completion based on specific characteristics may have consequences for how estimates are interpreted given known demographic variation in maltreatment behaviors. In addition, the examination of self-reported parenting behaviors considered problematic but below mandated reporting thresholds may provide insight into possible biases associated with self-reported child maltreatment behavior via IVR.

Caregiver Characteristics & Parenting Behaviors Associated with Child Maltreatment

Child maltreatment studies have typically assessed caregiver and child demographic characteristics as potential risk factors for child maltreatment (e.g., Brown, Cohen, Johnson, & Salzinger, 1998; Dubowitz et al., 2011; Mersky, Berger, Reynolds, & Gromoske, 2009). Caregiver demographics associated with risk for maltreatment include younger age, being female, being unmarried, having more children, and within lower socioeconomic status households (e.g., unemployed, in poverty, or less than high school education) (Sedlak et al., 2010; US DHHS, 2012). Children who are younger and female tend to be at the highest risk for child maltreatment (US DHHS, 2012). The relationships between maltreatment, race/ ethnicity and nativity continue to be subjects of debate; however, they remain a consistently measured risk factors given concerns about disproportionality and disparities in child welfare (Johnson-Motoyama, 2013).

More importantly, the examination of problematic parenting behaviors may help to identify caregivers who are more likely to engage in the types of maltreatment behaviors that would be addressed during an IVR portion of a survey. For instance, parenting stress has been identified consistently as a risk for child maltreatment, with higher levels associated with higher risk for maltreating behaviors, especially physically abusive behaviors (Hillson & Kuiper, 1994; Rodriguez & Green, 1997; Stith et al., 2009; Whipple & Webster-Stratton, 1991). Neglect is defined as a continuum of caregiver behaviors "that constitutes a failure to act in ways … necessary to meet the developmental needs of a child and which are the responsibility of a caregiver to provide" (Straus & Kantor, 2005, p. 20). Meeting a child's

basic physical needs (e.g., food and shelter) is critical for a child's general well-being and health (Casey et al., 2005), and meeting a child's basic supervisory needs (e.g., direct supervision and knowing whereabouts) helps to prevent accidental physical injury to a child (Landen, Bauer, & Kohn, 2003; Morrongiello, Klemencic, & Corbett, 2008). Therefore, a caregiver's inability to meet basic physical and supervisory needs may also indicate potential for maltreating behaviors (Magura & Moses, 1986; Straus & Kantor, 2005; Zuravin, 1991).

Aims of the Study

To our knowledge, our study is the first to assess the utility of IVR methods with parent self-report of child maltreatment behaviors. The use of IVR technology to capture actual abusive and neglectful parenting behaviors can advance our understanding of the scope of abuse and neglect experienced by children, especially for populations overlooked by current surveillance systems (Hammond, 2003). Concerns remain about whether the use of this technology results in significant differences between those who choose to complete the survey and those who dropout before completing the survey. Our aim is to describe survey completion behavior for caregivers surveyed with IVR, the biases that may arise from differences between those who complete and do not complete the survey, and whether these biases may affect our ability to generalize results of the survey to the population sampled. The study assesses whether survey completion is associated with a systematic bias in caregiver reports of child maltreatment relative to demographic characteristics and problematic parenting behaviors that were reported to a live interviewer during the CATI portion of the survey prior to transfer to the IVR portion.

Methods

Survey Design & Sample

The data used for this study come from a general population telephone survey conducted from March to October 2009 of 3,023 parents or legal guardians with children 12 years or younger residing in 50 cities in California. The survey employed a purposive geographic sample of 50 mid-sized cities (i.e., population between 50,000 and 500,000) randomly selected from 138 incorporated cities in California that were not adjacent to any other city in the sample. We then used list-assisted sampling to create a sampling frame of potential respondents. The listed sample was composed of addresses and telephone numbers obtained from a third party vendor who has access to these data from sources such as credit bureaus, credit card companies, utility company lists, and other companies that maintain lists. These lists were supplemented with any vendor lists of households with a child under the age of thirteen and then de-duplicated against each other before being randomized. List-assisted sampling combines random digit dialing with vendor-acquired listings in order to more effectively target sampling areas within a geographic area, such as are needed for the current study design (Gruenewald, Remer, & LaScala, 2014). When compared to traditional RDD techniques, listed samples are relatively unbiased, not highly correlated with socioeconomic status, and can be mitigated with the use of post-stratification weighting procedures (Brick, Waksber, Kulp, & Starer, 1995; Boyle, Bucuvalas, Piekarski, & Weiss, 2009; Kempf & Remington, 2007; Tucker, Lepkowski, & Piekarshi, 2002). All potential respondents were

sent a letter describing the study, informing them they may be contacted, and providing them an opportunity to opt out of the study by calling a toll-free telephone number.

A household was considered eligible for inclusion in the study if it contained at least one child twelve years old or younger who resided in the home at least 50% of the time, was an English- or Spanish-speaking household, and was located within one of the 50 selected cities. Respondents had to be age 18 years or older and a parent or legal guardian of the child and were chosen using a random selection procedure when more than one eligible respondent resided (i.e., two parents) in the household. Individuals who lived in institutional settings, who were not well enough to complete the interview, or did not speak English or Spanish were excluded from the study. The response rate for the survey was 47.4% (Freisthler & Gruenewald, 2013).

The final sample included 3,023 parents or legal guardians with children 12 years or younger with approximately 60 respondents per city (range of 47 to 74). We used post-stratification adjustments to increase generalizability to all 138 incorporated, mid-sized cities in California identified in the city-level sampling frame. Using a strategy similar to Brick and Kalton (1996), we weighted the study sample at the individual level using a single weight calculated from gender, race/ethnicity, and household type (i.e., single mother, single father, or two-parent household) to reflect the population attributes of these cities. Table 1 details the weighted descriptive characteristics of the full sample.

Respondents received $25 for participating in the 30 minute survey. The majority of the 25 minute survey was conducted with a live interviewer using computer-assisted telephone interviewing (CATI), which involved the interviewer sitting in front of the computer screen, the computer calling the respondent's telephone number, and the interviewer reading the survey from the computer screen and recording responses directly into the computer. In the CATI portion of the survey, parents/legal guardians were asked to self-report demographic information for themselves, the household, and a focal child who had the most recent birthday. They also self-reported parenting behaviors that did not require reporting to child protective services for the selected focal child but would still be considered problematic such as not providing healthy foods or a warm shelter and not safely monitoring a child under his/her care.

Respondents were then transferred to the IVR section that consisted of a maximum of 21 age-specific questions that were computer-administered, taking about 5 minutes to complete on average. In order for respondents to self-administer responses, all respondents were required to have a touch-tone phone to complete this portion of the survey. The IVR section of the survey primarily focused on past year parenting behaviors that could result in reportable instances of physical abuse or neglect due to placing a child at risk for serious harm. All neglect items were selected from the Multidimensional Neglectful Behavior Scale (MNBS; Kantor, Holt, & Straus, 2003) using a 4 point Likert-type scale response option ranging from 1 (*Never*) to 4 (*Always*); developmentally specific items were asked for focal children ages 0 to 4 years, 5 to 9 years, and 10 to 12 years. All physical abuse items were selected from the Conflict Tactics Scale, Parent-Child Version (CTSPC; Straus et al., 1998) using categories for the number of times these behaviors occurred in the past year ranging

from 1 (*Never*) to 4 (*More than 10 times*); all respondents were asked to answer the same items regardless of focal child age. Responses were encrypted with only the research team (and not the survey firm) having the encryption key. Respondents gave informed consent verbally over the phone, after being provided with detailed information on the voluntary nature of the survey, description of the sensitive nature of the questions, information about mandating reporting laws, and an explanation that the IVR technology is used to protect confidentiality of responses about parenting practices by the interviewer.

Dependent Variable

The dependent variable measured whether a respondent completed the IVR section of the survey (completion = 1) or dropped out prior to completing the IVR section of the survey (completion = 0). We defined survey dropout as a respondent having consecutive missing values for the last two questions of the IVR section. Responses that required the caregiver to input a number in the telephone, including "Don't Know" or "Refused" were not considered missing given the respondent was still engaged with the computerized interviewing system. This definition results in caregivers completing the survey responding to a minimum of 95% of the IVR questions, all of which we deemed crucial for estimating maltreatment behavior. All 3,023 respondents completed the CATI portion of the survey. A total of 2,812 respondents (93%) completed the IVR portion of the survey and 211 respondents (7%) dropped out during the IVR portion of the survey.

Independent Variables

Parenting Stress—Two self-report parenting stress items from the Dimensions of Discipline Inventory (DDI; Straus & Fauchier, 2011) were obtained in the CATI portion of the survey. The items measured caregiver self-report of behaviors related to feelings of stress and anger when his/her child misbehaved. A 4 point Likert-type scale was used for each item ranging from 1 "Never" to 4 "Always." We calculated a stress scale by taking the mean of both questions. Scores ranged from 1 (Low Stress) to 4 (High Stress) with a mean value of 1.9 and a standard deviation of 0.7. For this sample, the scale demonstrated moderate levels of internal consistency ($\alpha = 0.67$). These findings are consistent with psychometric properties of the original scale (M = 1.9; SD = 0.7; $\alpha = 0.64$) (Straus & Fauchier, 2011).

Physical & Supervisory Needs—The MNBS measures a range of behaviors associated with a caregiver's ability to meet a focal child's basic needs in the past year (Kantor et al., 2003). The survey uses a 4 point Likert-type scale ranging from 1 *Never* to 4 *Always*. Lower scores are indicative of fewer needs met and thus more neglectful behaviors. In the CATI portion of the survey, items were asked from the MNBS that did not necessitate reporting to child protective services due to the lower likelihood of physical harm to a child. However, these behaviors still may indicate risk for neglect given a potential failure to provide basic child needs (Straus & Kantor, 2005). Since the scale items from the MNBS were split between the CATI and IVR sections of the survey, the ability to construct reliable scales for this study was limited given CATI items are being used to assess potential bias in IVR responses. We conducted factor analysis and reliability tests and determined that the CATI items for physical needs and supervisory needs would be better used as separate single-item

constructs since there was low internal consistency across CATI-only items for physical needs ($\alpha = 0.291$) and supervisory needs ($\alpha = 0.294$). All single-item variables were computed as categorical variables (0 = *Never/Sometimes*, 1 = *Often/Always*). Evidence of construct validity for these items was determined by: (a) consistent significant associations observed between CATI-obtained physical and supervisory need items and IVR-obtained MNBS neglect items, and (b) limited to no significant associations observed between CATI-obtained physical and supervisory need items and IVR-obtained physical assault items from the CTSPC (results available upon request).

Physical Needs: Parenting behaviors categorized as meeting physical needs of a child were captured during the CATI portion of the survey by three items from the MNBS that were used to create two single-item constructs: a) provision of warm shelter and b) provision of healthy food. Provision of warm shelter was measured using one question for all age groups ("how often was the house warm enough when it was cold outside?"). Provision of healthy food was measured using questions regarding food variety specific to children ages 0 to 4 years ("how often did you provide your child with a variety of foods?") and children ages 5 to 12 years ("how often did you encourage your child to eat vegetables, fruit, and milk?").

Supervisory Needs: Parenting behaviors categorized as meeting supervisory needs of a child were measured during the CATI portion of the survey by three single-item constructs: a) safe monitoring of child's behavior, b) knowledge of child's location, and c) attention to misbehavior. "Safe monitoring of child's behavior" was created from items specific to children ages 0 to 4 years ("how often could you always hear your child when s/he cries and you are out of the room?"), 5 to 9 years ("how often did you NOT know where your child was playing when s/he was outdoors?" [reverse coded]), and 10 to 12 years ("how often did you call your child from work to check up on him/her?" [included *Not Applicable* option]). "Knowledge of a child's location" was created from items specific to children ages 0 to 4 years ("how often did you feel comfortable with the person that you left your child with?"), 5 to 9 years ("how often did you NOT know what your child was doing when s/he was not home? [reverse coded]), and 10 to 12 years ("how often have you known where you child was going after school?" [included *Not Applicable* option]). "Attention to misbehavior" was created from items specific to children ages 0 to 4 years ("how often did you distract your child when s/he tries to do something that could be unsafe like pull on electric plug or touch the stove?") and 5 to 12 years ("how often have you NOT cared if your child got in trouble at school?" [reverse coded & *Not Applicable-My child does not get into trouble* option]).

Demographic Characteristics—Caregivers reported age in years, gender (male/female), partnership status (married or cohabiting compared to single, divorced or widowed), unemployment status (unemployed/not unemployed), education completed (less than high school/high school diploma or more), preferred language spoken (English/Spanish), nativity (as defined by whether they were born in the U.S. or elsewhere), and number of children 12 years or younger to a live interviewer during the CATI portion of the survey. Caregivers also reported race/ethnicity (white, Hispanic, non-Hispanic black, non-Hispanic Asian, multi-racial, or other) during the CATI portion of the survey. Race/ethnicity was determined using multiple questions that asked respondents to report up to two racial/ethnic groups that

best describes the respondent's family of origin; any respondent that identified more than one racial/ethnic group was categorized as multi-racial. Each respondent were also asked to select one of eight categories that best described the total household income before taxes for the 2008 tax return, which ranged from less than \$10,000 to more than \$150,000 during the CATI portion of the survey. Income was kept as a categorical variable to indicate whether the household income was "\$40,000 or less" or "more than \$40,000" to capture low-income households that meet criteria for eligibility requirements to receive benefits from California programs, such as WIC, Food Stamps, and CalWORKS (185% of the Federal Poverty Limit; US DHHS, 2008).

Caregivers reported child demographic characteristics for child age in years and child gender (male/female) in both the CATI and IVR section of the survey; due to a large percent of dropouts occurring during the transition to IVR, child demographics reported to a live interviewer in the CATI portion of the survey were used in the analyses. Child age was categorized into three groups that parallel the age breakdown used by the Multidimensional Neglect Behavior Scale (MNBS)—0 to 4 years, 5 to 9 years, and 10 to 12 years (Kantor et al., 2003). Table 2 details each demographic variable with associated descriptive statistics.

Statistical Analyses

Trends in Survey Completion by Item—To assess general trends in dropout behavior, we tracked completion by item for the IVR section of the survey and graphed percentage of respondents completing each item (see Figure 1). Because age-specific items were included from the MNBS, trends are split up across focal children ages 0 to 4 years and 5 to 12 years.

Bivariate—We used chi square and *t* tests to compare whether respondents' completion of the IVR section of the survey was associated with respondent demographics, focal child demographics, and parenting behaviors using SPSS 21 (IBM Corp, 2012). Table 2 provides the results of the bivariate analyses using the analytic sample.

Multivariate—We used a multi-level model to assess which of these variables were associated with survey completion since the study design results in respondents (Level 1) being nested within cities (Level 2). We used the general form of the multilevel model:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_p X_p + e \quad \text{Level 1}$$

$$b_0 = g_{00} + u_0 \quad \text{Level 2}$$

For Level 1, Y was a binary outcome indicating whether or not a respondent completed the survey, measured at the person level. The variable, b_0, was the city-specific intercept. Variables, $b_{1\ to\ p}$, are the regression coefficients expressing the associations between p person-level predictors for demographic and parenting variables and the outcome of survey completion. The individual-specific residual or error is represented by the variable, e. For Level 2, $g0_0$ indicates the overall sample intercept for the equation predicting city-specific intercepts, and u_0 indicates the random city-specific residual component. At the highest level of analysis (Level 2, city level), we used only a constant to account for city-level clustering

that may be impact completion due to variation in social environments (Groves & Couper, 1998).

We used a unit-specific Bernoulli Hierarchical Generalized Linear Models (HGLM) with a logit link function to analyze the data (Raudenbush & Bryk, 2002). Analyses were conducted separately for (a) the full sample, (b) focal children ages 0 to 4, and (c) focal children 5 to 12. We used the HGLM module of the HLM Version 7 software (Raudenbush, Bryk, Cheong, Congdon, & de Toit, 2011). Education level was excluded from the final adjusted model due to multicollinearity in the model with poverty, preferred language, and nativity.

Missing Data—Cases with missing data were excluded from final analyses resulting in 220 cases (7%) being removed from the full model. Age-specific models were created from this analytic sample (e.g., 0 to 4 years = 77 missing cases; 5 to 12 years = 143 missing cases). Table 1 and table 2 show the weighted univariate statistics for both the full sample and analytic sample. We examined the effect of the missing values by conducting bivariate analyses (either chi-square or t-tests) comparing respondents with missing data with respondents without missing data. Overall, no statistically significant differences were observed between respondents with and without missing data by IVR completion status (χ^2 (1, n = 3023) = 0.35, p = 0.56). Household income was the single variable with the largest number of missing values (n = 115); missing income values were also independent of the dependent variable (χ^2 (1, n = 3023) = 0.81, p = 0.37).

Results

Of the 211 respondents (7%) who dropped out during the IVR section, 125 respondents dropped out during the transition from CATI to IVR (4% of total sample), which is defined as respondents leaving the survey (i.e. hanging up the phone) after the IVR section is initiated and prior to successfully completing the first non-demographic survey item in the IVR section. Reasons for dropping out during the transition from CATI to IVR included refusal to complete IVR at time of transition (n= 46), no touch-tone phone (n= 21), and unsuccessful recall to complete the survey after respondent hung-up during transition to IVR (n= 58). The remaining 86 respondents (3% of total sample) dropped out at various points during the IVR section. Figure 1 shows percent of respondents completing the IVR portion of the survey by item for focal children 0 to 4 years and for focal children 5 to 12 years. While reasons for discontinuing the IVR section were not specifically assessed for the remaining 86 respondents, trends in Figure 1 show two of the largest single-item drops in completion were for (a) the shift from the MNBS 4-point Likert response options to the CTS-PC frequency of behaviors response items (0.4% of total sample, n = 13) and (b) when respondents were asked to self-report "In the past year, how often have you hit [focal child] on some other part of the body besides the bottom with something like a belt, hairbrush, a stick or some other hard object?" (1.3% of the total sample, n = 39).

Table 1 shows the results of bivariate analyses performed to examine the relationship between IVR completion by respondents and selected demographic characteristics. Completion of the IVR section was independent of key caregiver demographic traits, such as

age, gender, race/ethnicity, unemployment status, and number of children 12 years or younger. Caregivers were more likely to complete the IVR portion of the survey if they were married or cohabitating, completed high school or more, preferred to complete the survey in English, were born in the United States, or reported a yearly household income of more than $40,000. Regarding focal child demographic characteristics, neither age nor gender was significantly different between groups. Table 2 shows the results of bivariate analyses performed to examine the relationship between IVR completion by respondents and selected parenting behaviors. No parenting behaviors were significantly different between respondents who completed and did not complete the IVR section of the survey.

Table 3 shows the results of the multilevel Bernoulli regression for the full sample. No parenting variables significantly differed between respondents who completed and did not complete the IVR section of the survey for the full sample. Of all demographic characteristics, only respondents born in the United States significantly differed between groups. U.S. born respondents were 2.25 times more likely to complete the survey than respondents born elsewhere in the full model. No other caregiver demographic characteristics or child demographic characteristics were significantly related to IVR completion.

The age-specific multilevel models included all demographic variables excluding child age and all parenting variables. Demographic behaviors did not generally differ between respondents who completed and did not complete the IVR portion of the survey. There were no statistically significant differences by caregiver or child demographic characteristics between respondents who completed and did not complete the IVR section of the survey for respondents who reported behaviors towards a focal child age 0 to 4 years (n = 870). Only nativity significantly differed by survey completion behavior for respondents who reported on a focal child age 5 to 12 years (n = 1933). After controlling for other demographic characteristics and parenting behaviors, those caregivers who were born in the United States were more likely to complete the survey than respondents born elsewhere (OR = 2.49, 95% CI = [1.24, 5.02].

For focal children 5 to 12 years (n = 1933), parenting stress, provision of physical needs, and provision of supervisory needs did not significantly differ between respondents who completed the survey and respondents who dropped out of the survey. Respondents who reported on a focal child ages 0 to 4 years (n = 870) did not differ by parenting stress, adequate provision of physical needs, knowledge of child location, or attention to misbehavior. However, those respondents who self-reported higher levels of safe monitoring of a focal child ages 0 to 4 during the past year (i.e. could often/always hear child when s/he cries and respondent is out of the room) were 2.97 times more likely to complete the survey than respondents who self-reported lower levels of safe monitoring (OR = 2.97, 95% CI = [1.18, 7.51]).

Conclusions

Survey completion behavior for the IVR portion of the survey reflected patterns indicated by previous literature assessing IVR use: initial and continuous dropout across the IVR survey

items (Galesic et al., 2006; Kreuter et al., 2008). However, the dropout rate of 7% is on the lower end of reported dropout rates for surveys using IVR technology (Galesic et al., 2006; Tourangeau et al., 2002). While trends in dropout by item suggested potential bias may arise due to differences between respondents who completed and did not complete the survey, the vast majority of demographic and parenting variables did not significantly differ by survey completion behavior.

For demographic characteristics, nativity stood out as a potential source of bias given that respondents born in the United States were more likely to complete the IVR portion of the survey than respondents born elsewhere, in both the full model and one age-specific model. Past research suggests that multiple factors influence respondent survey completion (Galesic et al., 2006; Tourangeau et al., 2002). While the survey was offered in both English and Spanish, it is possible that language barriers presented foreign-born respondents with a higher cognitive burden, contributing to respondent fatigue and survey drop-out (Bowling, 2005; Peytchev, 2009). While more research is necessary, it is also possible that foreign-born respondents with less familiarity and/or comfort with the use of IVR may have had concerns about reporting on sensitive topics such as child maltreatment, even when their confidentiality was assured.

Overall, respondents with parenting behaviors associated with maltreatment were no more likely to dropout during the IVR portion of the survey than other respondents in the full model and most age-specific models. The lone exception involved respondents who reported being more likely in the past year to often or always hear their young children (ages 0 to 4 years) when in another room when compared to their counterparts who reported they could never or sometimes hear their children when in another room. While more research is necessary, it is possible that the survey item could have influenced caregiver survey completion behavior among those who could not hear their child by increasing their awareness, thereby resulting in differential dropout behavior during the IVR portion (Feldman & Lynch, 1988). One implication of this observed difference in survey completion involves potentially lower estimates of supervisory neglect behaviors for caregivers with young children aged 0 to 4 years. Alternatively, this finding could have resulted by chance given the large number of comparisons conducted across the full sample and age-specific models.

Strengths and Limitations

The current study has several strengths. First, the study relied on a general population sample of caregivers to assess actual child maltreatment behaviors. Standard survey approaches have typically focused on measures of child abuse potential, perceived maltreatment by service providers, or victim recall of childhood experiences, which limits estimation of current rates of actual maltreatment behaviors (Wulczyn, 2009). The study also surveyed a large sample of caregivers to provide sufficient power for the large number of comparisons in this study. In addition, the study begins to address the need for innovative approaches to obtain more accurate estimates of child maltreatment behaviors. The study attempts to limit potential harm to respondents through the use of IVR methods in order to minimize socially desirable responding and mandated reporting requirements. Finally, the

use of mixed-survey modalities allowed for the use of responses to questions regarding parenting behaviors obtained via CATI procedures to assess differential dropout rates during the IVR section.

The study also has several limitations that need to be addressed in future studies. First, households without land lines were excluded from the survey, which likely resulted in the under-coverage of (a) households without phones, whose members are typically socioeconomically disadvantaged with limited health care access, and (b) cell-phone only households, whose members are typically younger in age, Hispanic, not married, and renters (Galesic et al., 2006). However, these limitations were balanced with a sampling design that provided an efficient approach to targeting families with children in the general population, and post-stratification adjustments were applied to correct for potential under-coverage. Another potential limitation stems from the IVR technology, which required respondents to have touch-tone phones to complete the IVR portion of the survey. This technology potentially biases IVR completion toward younger populations (Beach et al., 2010). However, we did not observe such bias in the bivariate and multivariate results, possibly due to the small percent of respondents (0.7%) who did not have touch-tone phones. Finally, respondents whose primary language was not English or Spanish (i.e., immigrant Asian populations) were potentially excluded from the survey, thereby limiting the generalizability of results to non-English and non-Spanish speaking populations.

With regard to measurement, we constructed the parenting stress variable using a validated construct from the DDI (Straus & Fauchier, 2011); however, the internal consistency for this scale is low due to the small number of items used to create this measure (Tavakol & Dennick, 2011). In addition, parenting items for physical and supervisory need were restricted to single-item constructs for this study due to the focus on non-reportable items answered in the CATI portion of the survey. However, the use of a single-item was preferable to a multi-item approach given the low internal consistency across items available in the CATI portion of the survey for these constructs. Future studies should consider the use of more robust indicators. Other risk factors associated with severe abuse or neglect behaviors, such as caregiver mental health, also were not included in the current models (e.g., Brown et al., 1998).

Finally, the current study is unable to distinguish if differences in groups are due to respondent reaction to the survey modality, survey fatigue, respondent comprehension of items, or respondent reaction to questions of a sensitive nature. More research is needed to determine the potential implications of demographic differences biasing child maltreatment estimates. For example, the significance of nativity for IVR may be due a variety of unmeasured factors, such as cultural differences related to response to an automated system, cognitive burden associated with answering items in a non-preferred language, or differential discomfort in responding to sensitive parenting questions. Moreover, future studies can improve upon the current study design by randomly assigning participants to complete CATI and IVR sections. This approach would allow direct comparison of dropout behavior by modality and response to sensitive items.

Significance

Innovative survey methodologies are needed to obtain more accurate estimates of child maltreatment, which are essential for better defining the scope of abuse and neglect and expanding prevention efforts to populations overlooked by current surveillance systems (Hamond, 2003). This study suggests that IVR methods may be appropriate to capture child maltreatment behaviors within the general population in a way that minimizes potential bias in self-reporting. Overall, survey completion behavior does not seem to be associated with a systematic bias related to parenting behaviors with the possible exception of safe monitoring behaviors for caregivers of children ages 0 to 4 years. While demographic differences were not generally associated with survey completion behavior, further research is needed to assess the appropriateness of this technology for use with populations that are diverse with regard to factors associated with nativity, such as language use and cultural orientation. Child welfare practitioners and researchers can benefit from further exploring such methodological innovations to obtain more accurate estimates of child maltreatment behaviors in the general population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Annual Update of the HHS Poverty Guidelines. 73 Fed. Reg. 3971. 2008 Jan 18.

Beach SR, Schulz R, Degenholtz HB, Castle NG, Rosen J, Fox AR, Morycz RK. Using audio computer-assisted self-interviewing and interactive voice response to measure elder mistreatment in older adults: Feasibility and effects on prevalence estimates. J Off Stat. 2010; 26(3):507–533. [PubMed: 21113391]

Bosnjak M, Tuten TL. Classifying response behavior in web-based surveys. Journal of Computer-Mediated Communication. 2001; 6(3) 0.

Bowling A. Mode of questionnaire administration can have serious effects on data quality. Journal of Public Health. 2005; 27(3):281–291. [PubMed: 15870099]

Boyle J, Bucuvalas M, Piekarski L, Weiss A. Zero Banks Coverage Error and Bias in Rdd Samples Based on Hundred Banks with Listed Numbers. Public Opinion Quarterly. 2009; 73(4):729–750.

Brick JM, Kalton G. Handling missing data in survey research. Statistical Methods in Medical Research. 1996; 5(3):215–238. [PubMed: 8931194]

Brick JM, Waksberg J, Kulp D, Starer A. Bias in list-assisted telephone samples. Public Opinion Quarterly. 1995; 59(2):218–235.

Brown J, Cohen P, Johnson JG, Salzinger S. A longitudinal analysis of risk factors for child maltreatment: Findings of a 17-Year prospective study of officially recorded and self-reported child abuse and neglect. Child Abuse & Neglect. 1998; 22(11):1065–1078. [PubMed: 9827312]

Casey PH, Szeto KL, Robbins JM, Stuff JE, Connell C, Gossett JM, Simpson PM. Child health-related quality of life and household food security. Archives of Pediatrics & Adolescent Medicine. 2005; 159(1):51–56. [PubMed: 15630058]

Child Trends. Child maltreatment. 2012. Available at: http://www.childtrends.org/?indicators=child-maltreatment

Cicchetti D, Toth SL. Child maltreatment. Annual Review of Clinical Psychology. 2005; 1:409–438.

Corkrey R, Parkinson L. Interactive voice response: Review of studies 1989-2000. Behavior Research Methods, Instruments, & Computers. 2002a; 34:342–353.

Corkrey R, Parkinson L. A comparison of four computer-based telephone interviewing methods: Getting answers to sensitive questions. Behavior Research Methods, Instruments, & Computers. 2002b; 34:354–363.

Couper MP, Groves RM. Household-level determinants of survey non-response. New Directions for Evaluation. 1996; 70:63–79.

Drake B, Jonson-Reid M. Some thoughts on the increasing use of administrative data in child maltreatment research. Child Maltreatment. 1999; 4:308–314.

Dubowitz H, Kim J, Black MM, Weisbart C, Semiatin J, Magder LS. Identifying children at high risk for a child maltreatment report. Child Abuse & Neglect. 2011; 35:96–104. [PubMed: 21376396]

Feldman JM, Lynch JG. Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. Journal of Applied Psychology. 1988; 73(3):421–435.

Finkelhor D, Turner H, Ormrod R, Hamby SL. Violence, Abuse, and Crime Exposure in a National Sample of Children and Youth. Pediatrics. 2009; 124:1411–1423. [PubMed: 19805459]

Freistler B, Gruenewald PJ. Where the individual meets the ecological: a study of parent drinking patterns, alcohol outlets, and child physicalabuse. Alcoholism: Clinical Experimental Research. 2013; 37(6):993–1000. http://dx.doi.org/10.1111/acer.12059.

Galesic M, Tourangeau R, Couper MP. Complementing random-digit-dial telephone surveys with other approaches to collecting sensitive data. American Journal of Preventive Medicine. 2006; 31(5):437–443. [PubMed: 17046416]

Groves, RM.; Couper, MP. Nonresponse in Household Interview Surveys. New York: Wiley and Sons, Inc; 1998. ISBN 0-471-18245-1

Groves RM, Cialdini RB, Couper MP. Understanding the decision to participate in a survey. Public Opinion Quarterly. 1992; 56(4):475–495.

Gruenewald, PJ.; Remer, L.; LaScala, EA. Testing a social ecological model of alcohol use: The California 50-city study. Addiction. 2014. http://dx.doi.org/10.1111/add.12438

Hammond WR. Public health and child maltreatment prevention: the role of the centers for disease control and prevention. Child Maltreatment. 2003; 8(2):81–83. [PubMed: 12735710]

Hardt J, Rutter M. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. Journal of Child Psychology and Psychiatry. 2004; 45:260–273. [PubMed: 14982240]

Hillson JMC, Kuiper NA. A stress and coping model of child maltreatment. Clinical Psychology Review. 1994; 14(4):261–285.

Hussey JM, Chang JJ, Kotch JB. Child maltreatment in the United States: prevalence, risk factors, and adolescent health consequences. Pediatrics. 2006; 118:933–942. [PubMed: 16950983]

IBM Corp. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp; 2012.

Johnson-Motoyama, M. Does a paradox exist in child well-being risks among foreign-born Latinos, U.S.-born Latinos, and whites? Findings from 50California cities. Child Abuse & Neglect. 2013. http://dx.doi.org/10.1016/j.chiabu.2013.09.011

Kantor, GK.; Holt, M.; Straus, MA. The parent-report multidimensional neglectful behavior scale. Durham, NH: Family Research Laboratory; 2003.

Kempf AM, Remington PL. New challenges for telephone survey research in the twenty-first century. Annu. Rev. Public Health. 2007; 28:113–126. [PubMed: 17094769]

Kreuter F, Presser S, Tourangeau R. Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. Public Opinion Quarterly. 2008; 72(5):847–865.

Landen MG, Bauer U, Kohn M. Inadequate supervision as a cause of injury deaths among young children in Alaska and Louisiana. Pediatrics. 2003; 111(2):328–331. [PubMed: 12563059]

Lesser VM, Newton LA, Yang D. Comparing item nonresponse across different delivery modes in general population surveys. Survey Practice. 2012; 5(2)

Magura, S.; Moses, BS. Outcome measures for child welfare services: Theory and applications. Washington, DC: Child Welfare League of America; 1986.

Mersky JP, Berger LM, Reynolds AJ, Gromoske AN. Risk factors for child and adolescent maltreatment: A longitudinal investigation of a cohort of inner-city youth. Child Maltreatment. 2009; 14(1):73–88. [PubMed: 18596199]

Midanik LT, Greenfield TK. Interactive voice response versus computer-assisted telephone interviewing (CATI) surveys and sensitive questions: The 2005 National Alcohol Survey. Journal of Studies on Alcohol and Drugs. 2008; 69(4):580–588. [PubMed: 18612574]

Morrongiello BA, Klemencic N, Corbett M. Interactions between child behavior patterns and parent supervision: Implications for children's risk of unintentional injury. Child Development. 2008; 79(3):679–638.

Peytchev A. Survey breakoff. Public Opinion Quarterly. 2009; 73(1):74–97.

Raudenbush SW, Bryk AS, Cheong YF, Congdon RT, du Toit M. HLM 7. Lincolnwood, IL: Scientific Software International. 2011

Raudenbush, SW.; Bryk, AS. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd Ed.. Thousand Oaks, CA: Sage Publications, Inc.; 2002.

Rodriguez CM, Green AJ. Parenting stress and anger expression as predictors of child abuse potential. Child Abuse & Neglect. 1997; 21(4):367–377. [PubMed: 9134265]

Rodriguez H, von Glahn T, Rogers WH, Chang H, Fanjiang G, Safran DG. Evaluating patient's experiences with individual physicians: A randomized trial of mail, internet, and interactive voice response telephone administration of surveys. Medical Care. 2006; 44(2):167–174. [PubMed: 16434916]

Schroder KEE, Johnson CJ, Wiebe JS. Interactive voice response technology applied to sexual behavior self-reports: A comparison of three methods. AIDS Behavior. 2007; 11:313–323. [PubMed: 16841191]

Sedlak, AJ.; Ellis, RT. Trends in Child Abuse Reporting. In: Korbin, JE.; Krugman, RD., editors. Handbook of Child Maltreatment. New York: Springer; 2014.

Sedlak, AJ.; Mettenburg, J.; Basena, M.; Petta, I.; McPherson, K.; Greene, A.; Li, S. Fourth National Incidence Study of Child Abuse and Neglect (NIS–4): Report to Congress. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families; 2010. Available from http://www.acf.hhs.gov/sites/default/files/opre/nis4_report_congress_full_pdf_jan2010.pdf

Stith SM, Liu T, Davies LC, Boykin EL, Alder MC, Harris JM, Som A, McPherson M, Dees JEMEG. Risk factors in child maltreatment: A meta-analytic review of the literature. Aggression and Violent Behavior. 2009; 14:13–29.

Straus MA, Kantor GK. Definition and measurement of neglectful behavior: Some principles and guidelines. Child Abuse & Neglect. 2005; 29:19–29. [PubMed: 15664423]

Straus MA, Hamby SL, Finkelhor D, Moore DW, Runyan D. Identification of child maltreatment with the Parent-Child Conflict Tactics Scales: Development and psychometric data for a national sample of American parents. Child Abuse & Neglect. 1998; 22(4):249–270. [PubMed: 9589178]

Straus, MA.; Fauchier, A. Manual for the Dimensions of Discipline Inventory (DDI). Durham, NH: Family Research Laboratory, University of New Hampshire; 2011.

Tavakol M, Dennick R. Making sense of Cronbach's alpha. International Journal of Medical Education. 2011; 2:53–33.

Tourangeau R, Smith TW. Asking sensitive questions: The impact of data collection mode, question format, & question context. Public Opinion Quarterly. 1996; 60(2):275–304.

Tourangeau R, Steiger DM, Wilson D. Self-administered questions by telephone: Evaluating interactive voice response. Public Opinion Quarterly. 2002; 66(2):265–278.

Tucker C, Lepkowski JM, Piekarski L. The current efficiency of list-assisted telephone sampling designs. Public Opinion Quarterly. 2002; 66:321–338.

Turner, CF.; Forsyth, BH.; O'Reilly, J.; Cooley, PC.; Smith, TK.; Rogers, SM.; Miller, HG. Automated self-interviewing and the survey measurement of sensitive behaviors. In: Couper, Mick P.; Baler, Reginald P.; Bethlehem, Jelke; Cynthia Z, Clark; Martin, Jean; Nicholls, William, II;

O'Reilly, James M., editors. Computer-Assisted Survey Information Collection. New York: Wiley and Sons, Inc.; 1998. ISBN 0-471-17848-9

U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. Child Maltreatment 2011. 2012. Available from http://www.acf.hhs.gov/programs/cb/research-data-technology/statistics-research/child-maltreatment

Whipple EE, Webster-Stratton C. The role of parental stress in physically abusive families. Child Abuse & Neglect. 1991; 15:279–291. [PubMed: 2043979]

Wulczyn F. Epidemiological perspectives on maltreatment prevention. The Future of Children. 2009; 19(2):39–66. [PubMed: 19719022]

Zuravin SJ. Child neglect: A review of definitions and measurement research. Neglected children: Research, practice, and policy. 1999:24–46.
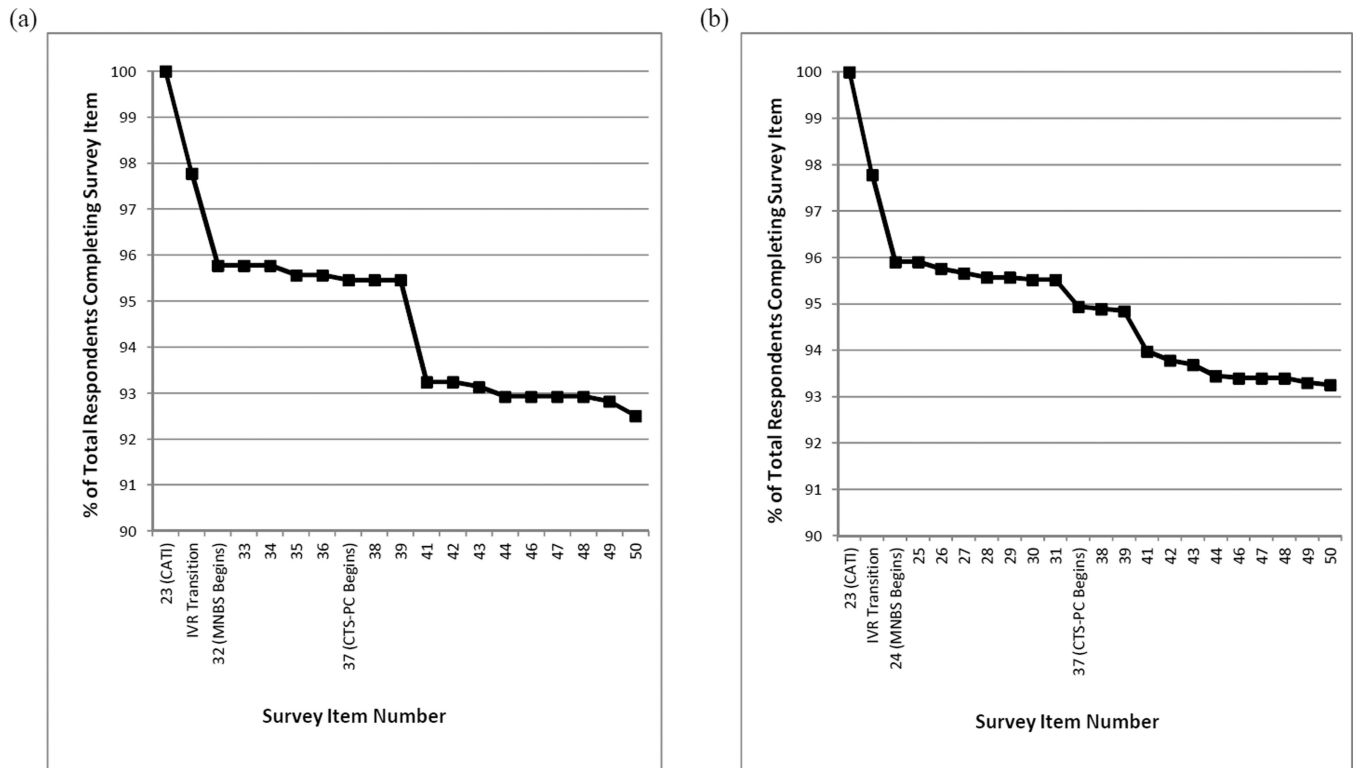
(a)

(b)

Figure 1.
Caregiver completion by IVR survey item for (a) focal child ages 0 to 4 years (n = 947) and
(b) focal child ages 5 to 12 years (n = 2076).

Table 1

Weighted Descriptive and Bivariate Results for Demographics Variables

	Full Sample (N = 3023)		Analytic Sample (n = 2803)		Complete (n = 2614)	Incomplete (n = 189)		
	N	Weighted % or Mean (SD)	n	Weighted % or Mean (SD)	Weighted % or Mean (SD)	Weighted % or Mean (SD)	χ^2/t	p
Caregiver								
Age (years)	3023	39.4 (8.4)	2803	39.4 (8.1)	39.4 (8.1)	38.9 (8.6)	0.83	0.409
Gender								
Male	1050	48.2	971	47.6	93.4	6.6	0.66	0.418
Female	1973	51.8	1832	52.4	92.7	7.3		
Race/Ethnicity								
White	1753	49.4	1641	50.0	94.1	5.9	7.80	0.168
Hispanic	733	30.6	683	30.3	91.1	8.9		
Non-Hispanic Black	111	4.8	105	4.6	92.1	7.9		
Non-Hispanic Asian	236	10.2	211	10.0	93.2	6.8		
Multi-racial	92	2.5	89	2.6	94.4	5.6		
Other	84	2.6	74	2.5	94.2	5.8		
Partnership Status								
Married/Co-Habit	2673	76.5	2477	77.2	93.8	6.2	7.79	0.005
Other	350	23.5	326	22.8	90.6	9.4		
Unemployment Status								
Unemployed	218	8.7	204	8.7	92.5	7.5	0.10	0.750
Other	2804	91.3	2599	91.3	93.1	6.9		
Education Completed								
< High School	150	6.2	131	6.0	85.1	14.9	17.25	< 0.001
High School	2873	93.8	2672	94.0	93.5	6.5		
Household Income								
$40,000 or Less	616	25.9	588	25.2	91.3	8.7	4.33	0.037
More than $40,000	2292	74.1	2215	74.8	93.6	6.4		
Preferred Language								

	Full Sample (N = 3023)		Analytic Sample (n = 2803)		Complete (n = 2614)	Incomplete (n = 189)		
	N	Weighted % or Mean (SD)	n	Weighted % or Mean (SD)	Weighted % or Mean (SD)	Weighted % or Mean (SD)	χ²/t	p
CATI in English	2817	92.2	2620	92.3	93.5	6.5		
CATI in Spanish	206	7.8	183	7.7	87.9	12.1	9.59	0.002
Nativity								
Born in the US	2306	73.6	2161	74.0	94.4	5.6		
Born outside of US	717	26.4	642	26.0	89.1	10.9	23.47	< 0.001
# of Children 12 yrs	3023	1.9 (0.8)	2803	1.9 (0.8)	1.9 (0.8)	1.8 (0.9)	1.67	0.094
Focal Child								
Child Age								
0 to 4 years	947	31.2	870	30.5	92.8	7.2		
5 to 9 years	1188	39.1	1111	39.2	92.0	8.0		
10 to 12 years	888	29.7	822	30.2	94.5	5.5	4.65	0.098
Child Gender								
Male	1565	51.0	1463	51.8	92.2	7.8		
Female	1454	49.0	1340	48.2	94.0	6.0	3.46	0.063

Table 2

Weighted Descriptive and Bivariate Results for Parenting Variables

	Full Sample (N = 3023)		Analytic Sample (n = 2803)		Complete (n = 2614)	Incomplete (n = 189)		
	N	Weighted % or Mean (SD)	n	Weighted % or Mean (SD)	Weighted % or Mean (SD)	Weighted % or Mean (SD)	χ^2/t	*p*
Parenting Stress	2984	2.0 (0.7)	2803	2.0 (0.7)	2.0 (0.6)	2.0 (0.8)	−0.92	0.360
Warm Shelter								
Never/Sometimes	137	5.2	120	5.0	92.0	8.0		
Often/Always	2875	94.8	2683	95.0	93.1	6.9	0.22	0.636
Healthy Food								
Never/Sometimes	130	4.3	117	4.2	93.2	6.8		
Often/Always	2871	95.7	2686	95.8	93.0	7.0	0.01	0.934
Knowledge of Location								
Never/Sometimes	107	4.2	98	4.1	88.6	11.4		
Often/Always	2874	95.8	2705	95.9	93.2	6.8	3.61	0.058
Safe Monitoring								
Never/Sometimes	442	16.8	407	16.4	92.3	7.7		
Often/Always	2236	73.0	2099	73.4	92.9	7.1		
Not Applicable	328	10.2	297	10.2	95.4	4.6	2.86	0.239
Attention to Misbehavior								
Never/Sometimes	265	10.1	247	10.3	91.6	8.4		
Often/Always	2240	74.0	2108	73.8	93.4	6.6		
Not Applicable	489	15.8	448	15.9	92.1	7.9	1.94	0.379

Table 3

Unit Specific Model of Respondent Characteristics Regressed on Survey Completion for Full Sample (n = 2803)

Variable	Model 1: Demographics		Model 2: Parenting		Model 3: Full	
	OR	CI	OR	CI	OR	CI
Intercept (City-Level Average)	3.91	[0.71, 21.58]	11.28	[2.24,56.89]	4.42	[0.44,44.84]
Level 1 (Individual):						
Caregiver Demographic Characteristics						
Age	1.00	[0.97, 1.04]			1.00	[0.97, 1.04]
Male	1.02	[0.64, 1.61]			1.03	[0.65, 1.63]
Race/Ethnicity (ref: White)						
Hispanic	0.90	[0.50, 1.60]			0.90	[0.51, 1.60]
Non-Hispanic Black	0.71	[0.27, 1.90]			0.69	[0.26, 1.84]
Non-Hispanic Asian	1.64	[0.63, 4.24]			1.70	[0.65, 4.43]
Multi-racial	1.00	[0.34, 2.94]			0.99	[0.34, 2.93]
Other	1.60	[0.41, 6.23]			1.55	[0.40, 6.00]
Currently married/cohabit	1.57	[0.86, 2.87]			1.52	[0.83, 2.78]
Unemployed	1.10	[0.53, 2.87]			1.09	[0.53, 2.26]
HH Income $40,000	1.10	[0.62, 1.97]			1.09	[0.61, 1.95]
English Speaking Preferred	0.98	[0.39, 2.43]			1.01	[0.40, 2.60]
US Born	2.26	[1.27, 4.04] **			2.25	[1.26, 4.01] **
Number of children <13 yrs	1.20	[0.93, 1.54]			1.20	[0.93, 1.54]
Child Demographic Characteristics						
Male	0.75	[0.50, 1.12]			0.74	[0.49, 1.11]
Age Group (ref: 0 to 4 yrs)						
5 to 9 years	0.94	[0.58, 1.53]			0.94	[0.55, 1.60]
10 to 12 years	1.40	[0.76, 2.59]			1.51	[0.74, 3.06]
Parenting Characteristics						
Parenting Stress			0.85	[0.63, 1.16]	0.82	[0.60, 1.13]
Warm Shelter (ref: Never/Sometimes)						

Variable	Model 1: Demographics		Model 2: Parenting		Model 3: Full	
	OR	CI	OR	CI	OR	CI
Often/Always			0.93	[0.37, 2.36]	0.73	[0.27, 2.00]
Healthy Food (ref: Never/Sometimes)						
Often/Always			0.85	[0.33, 2.19]	0.89	[0.31, 2.55]
Knowledge of Location (ref: Never/Sometimes)						
Often/Always			1.70	[0.71, 4.09]	1.37	[0.54, 3.50]
Safe Monitoring (ref: Never/Sometimes)						
Often/Always			1.07	[0.65, 1.78]	1.33	[0.70, 2.56]
Not Applicable			1.58	[0.64, 3.93]	1.36	[0.49, 3.78]
Attention to Misbehavior (ref: Never/Sometimes)						
Often/Always			1.30	[0.69, 2.42]	1.20	[0.63, 2.32]
Not Applicable			0.97	[0.46, 2.05]	0.87	[0.38, 1.99]
	T	(SE)	T	(SE)	T	(SE)
Variance Components (τ)	0.302	(0.119) ***	0.322	(0.122) ***	0.311	(0.121) ***

***
 p < .001,

**
 p < .01,

*
 p < .05