

Feasibility of Interactive Voice Response Methods in Health Behavior Research With Immigrant Latinos

Joseph G. Grzywacz, PhD; Sara A. Quandt, PhD; Antonio Marín, MA; Phillip Summers, MPH; Wei Lang, PhD; Thomas Mills, MS; Carlos Evia, PhD; Julia Rushing, MS; Thomas A. Arcury, PhD

Objectives: Determine the feasibility of using interactive voice response (IVR) for conducting daily diary research with immigrant Latinos. **Methods:** Baseline data containing demographic, health, and job-related characteristics were obtained from immigrant Latino construction workers (N=119). Participants also completed an IVR-based daily diary for 21 consecutive days. **Results:** Over one third (37%) of the sample adhered to the 21-day protocol; a comparable percent-

age (38.7%) initiated but did not adhere to the protocol; and 24% never began the diary protocol. Adherence was generally not predicted by demographic, health, or job-related characteristics. **Conclusions:** IVR technology makes diary studies with Latino samples feasible.

Key words: Latinos, interactive voice response, methods evaluation, daily diary

Am J Health Behav. 2013;37(2):238-247

DOI: <http://dx.doi.org/10.5993/AJHB.37.2.11>

Study designs that use repeated measures over discrete periods of time, like end-of-day diary studies, offer several strengths for understanding health behaviors.^{1,2} One key benefit is reduced measurement problems because of the shorter time between experiencing an event and subsequent reporting of it.¹ Another benefit of daily diary study designs is that they are able to capture “life as it is lived.”³⁻⁵ The repeated observations over a series of multiple days allow the capture of both highly routinized aspects of daily life and more unusual types of events. Greater capture of the circumstances and experiences of daily life allows for a more complete, and, presumably valid, description of human behavior and is, therefore, valuable for health research. Further, because

observations are captured in a sequence of time, daily diaries enable analytic techniques that allow maximal causal inference from observational data because investigators can capture and model the temporal ordering of observed phenomenon, thereby allowing individuals to essentially serve as their own controls.^{1,3}

Fielding daily diary studies can be challenging in low-literacy populations. Collecting diary data is challenging because keeping a diary typically requires participants to be able to read and potentially write, regardless of whether data are captured using paper and pencil or an electronic device (eg, personal digital assistants, tablet PC). Literacy issues can be partially remedied by the use of interviewer-administered methods (either in person or by telephone); however, interviewer-administered methods require substantial personnel time and may not be feasible because of either costs or sensitivity of the topic under study. Even more problematic is the coordination effort that may be needed to ensure that interviewers are available when data are to be collected. Irregular or chaotic daily schedules of study participants can compromise even the best-laid data collection plans.

Interactive voice response (IVR) technology offers substantial promise to health research, particularly research with low-literacy populations.⁶ IVR technology uses stored, prerecorded questions

Joseph G. Grzywacz, Antonio Marín, Phillip Summers, Thomas A. Arcury, Department of Family and Community Medicine, Center for Worker Health, Wake Forest School of Medicine, Winston-Salem, NC. Sara A. Quandt, PhD, Department of Epidemiology and Prevention, Center for Worker Health, Wake Forest School of Medicine, Winston-Salem, NC. Wei Lang, Julia Rushing, Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC. Thomas Mills, MS, Myers-Lawson School of Construction, Occupational Safety and Health Research Center, Virginia Polytechnic Institute and State University, Blacksburg, VA. Carlos Evia, PhD, Department of English, Occupational Safety and Health Research Center, Virginia Polytechnic Institute and State University, Blacksburg, VA.
Correspondence Dr Grzywacz; grzywacz@wakehealth.edu

tagged with voice extensible markup language (VoiceXML) to deliver audible questions to study participants over the telephone. Participants can respond to questions by entering a number as a text response on their telephone's keypad or by responding orally. This technology has many benefits. The use of a telephone and audio-based data collection system eliminates the need for participants to read or write, and it can accommodate non-English-speaking individuals. IVR ensures that all participants receive the same stimulus, thereby minimizing potential interviewer effects. The date and time of each diary entry are electronically stamped, allowing rigorous evaluation of adherence to the data collection protocol. Finally, provided they have a telephone, participants can access the data collection system virtually anytime and from anywhere.

IVR technology has been used in research for over 20 years.⁶ IVR has been used most extensively in studies of alcohol use and misuse. Collins and colleagues⁷ documented that IVR technology yielded reliable data from social drinkers, while having several benefits, such as ease of use by participants and date/time stamping of data entries. Midanik and Greenfield⁸ evaluated the use of IVR in the context of the 2005 National Alcohol Survey and reported that IVR was particularly useful for individuals for whom survey content was sensitive. Outside of alcohol research, IVR has been used to study sexual behavior,^{9,10} mistreatment of elders,¹¹ postpartum depression among low-income mothers,¹² postdischarge adverse events,¹³ and medication adherence.^{14,15} Absent from this growing area of research is any application of the method with low-literacy, non-English-speaking individuals. One study did compare the use of IVR in contrast to paper-and-pencil methods of collecting data on symptoms from Spanish-speaking patients and found that each method yielded similar results.¹⁶

The goal of this study was to develop a better understanding of the potential value of IVR technology in fielding daily diary studies in low-literacy, Spanish-speaking populations. In this study the focus was immigrant Latino workers in the residential construction industry. Latino construction workers provide an excellent model for evaluating the feasibility of IVR technology as they typically have modest levels of formal education,¹⁷ suggesting limited literacy in any language. The length and time of the workday can be highly variable and heavily influenced by external factors, such as weather-related delays and overall demand for services. Finally, immigrant Latino construction workers experience an elevated burden of occupational injuries, many of which are believed to go unreported either because they are perceived as minor or because workers are reluctant to report the injury to their employer out of fear of reprisal or sanction.¹⁸⁻²⁰

This analysis examines the feasibility of using IVR technology to field a daily diary study focused on work-related stressors and safety behavior in

the workplace among immigrant Latino residential construction workers. To achieve this objective, this analysis describes participant adherence to a 21-day end-of-day daily diary protocol using IVR technology; delineates demographic, individual, and work-related sources of nonadherence to the diary protocol; and documents variation in key call attributes (the number and duration of diary entries) and responses to a debriefing interview about the IVR technology by participant adherence status.

METHODS AND MATERIALS

Sample

Participants were a nonprobability sample (N=119) of residential construction workers who self-identified as Latino. Participants were recruited in partnership with HOLA (Hogar Latino) of Wilkes County, North Carolina. HOLA is a 501(c)(3) nonprofit organization that serves the Latino communities of Wilkes and surrounding counties in western North Carolina, a region of the country characterized as a "new settlement community" because of the sharp increase in Latino population since 1990.²¹ A main emphasis of HOLA's mission is to empower the Latino community to improve their health through promoting education and by strengthening understanding and cooperation between Latinos and the broader community. HOLA staff recruited residential construction workers through a combination of techniques including through known construction workers in existing social networks, snowball sampling (identification of individuals who may be study eligible by existing participants), and referral (identification of individuals who may be study eligible by nonparticipants). Study inclusion criteria were age 18 years or older, Latino (self or parents born in a Latin American country or self-identified as Latino or Hispanic), and employment for 35 or more hours per week in construction. There were no exclusion criteria.

The study design called for workers in each of 3 construction trades: framing, roofing, and general labor. Individuals were classified as being a framer if they were usually involved in building or assembling floors, walls, or roofs using lumber or light-gauge metal framing at least 3 days per week. Participants were classified as roofers if they were usually involved in applying roof shingles, removing existing roof shingles, applying hot roofing tar or built-up roofing, or carrying roofing materials onto a roof at least 3 days per week. Individuals who self-identified as being a general construction worker and did not meet the definition of framer or roofer were classified as being a general construction worker.

Data Collection

The data for this analysis come from 3 elements of the overall project. The first data collection element was a baseline interviewer-administered interview, which assessed stable attributes of the individual (eg, age, country of origin), occupational characteristics (eg, years in construction, primary

tasks performed in construction), and a health history (eg, presence of chronic conditions). These data were collected by 4 trained interviewers, all of whom were native Spanish-speakers. Interviewer training consisted of a thorough review of study purpose, screening and recruitment procedures, line-by-line review of the interviewer-administered questionnaires, and progressively more realistic practice interviews. The baseline interview was completed by N=119 participants and took an average of 48.1 (SD = 62.0) minutes to complete. Participants received a \$15 incentive.

The second data collection element was a 21-day daily diary implemented using IVR technology (PlumVoice, Cambridge MA). The 21-day period was selected based on practical considerations; specifically, we wanted to capture at least 2 weeks of workday observations while accommodating likely fluidity in work schedules (eg, missed work because of inclement weather). The IVR data collection system was designed and implemented using VoiceXML. The VoiceXML application guided participants through a series of questions about daily health symptoms and a variety of work-related experiences and exposures. The VoiceXML was programmed to follow a logic selection pattern dependent on the type of work done on each day. This logic provided flexibility such that individuals who did not work on a given day were not asked work-related questions, whereas other questions relevant to specific tasks (eg, ladder safety) could be skipped if deemed irrelevant (eg, reported not using a ladder during that work day). Participants received up to \$73 for completing the 21-day daily diary protocol, which essentially translates into \$3 per call plus \$10 to cover cell phone minutes that may have been used to participate in the study.

The third data collection element was a short debriefing interview implemented by a trained, native Spanish-speaking interviewer. The debriefing interview consisted of structured questions about participants' experiences with the IVR technology, as well as open-ended questions probing for detail when needed. This analysis draws on responses to the structured questions.

The content of each data collection element underwent thorough translation and back-translation procedures. Content from validated Spanish instruments (eg, SF-12 questionnaire for assessing health-related quality of life) were used without modification. English-only instruments and items developed for this project were translated into Spanish by a native Spanish speaker. All items were then back-translated into English by an English speaker. Discrepancies identified in the back translation were corrected through consensus and incorporated into both the Spanish and English versions of questionnaires.²²

Measures

Participants were classified into one of 3 mutu-

ally exclusive groups based upon their number of completed calls during the 21-day diary period. Participants were classified as noncompleters if they never completed a daily diary using the IVR system. Participants who logged into the IVR system at least once but completed fewer than 85% of the calls (fewer than 18 calls during the 21-day period) were classified as nonadherent. Adherent participants were those who completed a telephone diary on 85% or more of the call days (ie, 18 or more calls during the 21-day period).

Demographic characteristics. The baseline questionnaire contained several measures of basic demographic characteristics, by which adherence to the diary protocol might differ. These characteristics included *age* (measured in years), gender, marital status (single/unmarried, married/living as married with partner in the United States, and married/living as married with partner in country of origin), educational attainment based upon the grading system used in Mexico and other Latin American countries (ie, primary education or less, secondary education, and preparatory education or higher), and length of residence in the United States (≤ 7 years, 8-15 years, ≥ 16 years).

Health. Individuals with poor physical or mental health may have compromised ability to participate in a data collection project requiring completion of a daily questionnaire over an extended period (21 days) of time. Health was assessed with 4 instruments in the baseline questionnaire. The first instrument was the Disabilities of the Arm, Shoulder, and Hand (DASH) questionnaire, developed to measure impairment *and* limitations in activities due to disorders of the upper limbs. The DASH comprises 30 items and has been translated and validated in Spanish.²³ The second measure, the Roland-Morris Disability Questionnaire (RDQ), is a 24-item instrument designed to measure physical disability due to low back pain. The RDQ has been translated into Spanish²⁴ and is reported to have good internal consistency.²⁵ Third, the Medical Outcomes Study Short Form-12 Health Survey, V 2.0, Acute form²⁶ was used to assess health-related quality of life; proprietary algorithms were used to construct norms-based scores for physical and mental health (ie, PCS and MCS). Finally, depressive symptoms were assessed with the Center for Epidemiological Studies-Depression scale (CES-D).²⁷ The CES-D is a widely used instrument, with previous studies demonstrating adequate internal consistency (.81 - .90) and construct validity (.80) in Mexican and Mexican American samples.²⁸⁻³¹ The 20 items of the CES-D are summed, with higher values indicating greater depressive symptoms.

Occupational characteristics. Discrete features of participants' jobs may influence their ability to adhere to the diary protocol. Individuals were classified at recruitment based on the inclusion criterion into one of 3 mutually exclusive categories (ie, framer, roofer, and general construction).

Table 1
Demographic Characteristics by Adherence Status to the 21-day
IVR Telephone Diary

Characteristics	Sample N (%)	Noncompleters N (%)	Nonadherent N (%)	Adherent N (%)	p-value
Sample	119 (100)	29 (24.4)	46 (38.7)	44(37.0)	
Period					.04
1 (9/24/10 to 10/7/10)	60 (51.3)	8 (13.3)	26 (43.3)	26 (43.3)	
2 (10/8/10 to 11/1/10)	57 (48.7)	19 (33.3)	20 (35.1)	18 (31.6)	
Age					.50
19 to 28	42 (35.3)	7 (16.7)	20 (47.6)	15 (35.7)	
29 to 38	56 (47.1)	17 (30.4)	18 (32.1)	21 (37.5)	
≥ 39	21 (17.6)	5 (23.8)	8 (38.1)	8 (38.1)	
Marital Status					.86
Not married	43 (36.1)	13 (30.2)	15 (34.9)	15 (34.9)	
Married, unaccompanied	12 (10.1)	2 (16.7)	5 (41.7)	5 (41.7)	
Married, accompanied	64 (53.8)	14 (21.9)	26 (40.6)	24 (37.5)	
Educational Attainment					.74
Primary or less	44 (37.0)	8 (18.2)	20 (45.5)	16 (36.4)	
Secondary	50 (42.0)	14 (28.0)	17 (34.0)	19 (38.0)	
Preparatory or more	25 (21.0)	7 (28.0)	9 (36.0)	9 (36.0)	
Years in the US					.66
≤ 7 years	44 (37)	13 (29.5)	15 (34.1)	16 (36.4)	
8 to 15 years	58 (48.7)	13 (22.4)	22 (37.9)	23 (39.7)	
≥ 16 years	17 (14.3)	3 (17.6)	9 (52.9)	5 (29.4)	
Country of Origin					.16
Mexico	91 (77.0)	26 (28.6)	33 (36.3)	32 (35.2)	
Other	28 (23.0)	3 (10.7)	13 (46.4)	12(42.9)	

Note.
p-values obtained using chi-square test

Hours worked per week in construction and weeks worked in construction in the previous 12 months were reported by participants during the baseline

interview.

Individuals in more punitive or demanding types of work environments may be less able or interest-

Table 2
Baseline Health Characteristics by Adherence Status to the 21-Day IVR Telephone Diary

Characteristics	Sample M (SD)	Noncompleters M (SD)	Nonadherent M (SD)	Adherent M (SD)	p-value
Physical Health					
DASH score	3.4 (4.9)	3.5 (4.8)	4.2 (6)	2.6 (3.5)	.29
Roland-Morris score	0.3 (1.7)	0.4 (2.4)	0.1 (0.7)	0.5 (1.9)	.54
PCS	46.4 (5.0)	46.5 (5.2)	45.7 (5.7)	47 (4.1)	.54
Mental Health					
CES-D score	12.9 (4.8)	13.3 (4.9)	12.9 (5)	12.7 (4.7)	.84
MCS	40.9 (8.0)	41.3 (9)	40.5 (8.2)	41 (7.5)	.92

Note.
 p-values obtained using a one-way ANOVA

ed in participating in a research study with repeated and ongoing requirements. Several variables reflecting the organization of work were measured in the baseline questionnaire using established instruments. Job design and performance variables were assessed using a modified version of the Job Content Questionnaire (JCQ).³² Based on previous research with immigrant Latinos suggesting difficulty responding to affectively based response items (eg, strongly agree versus strongly disagree),^{33,34} the JCQ was modified to use a 4-point frequency-based response set (never to always). Control was assessed with 3 items tapping opportunities to exert control over work (eg, "How often are you allowed to make your own decisions about your work?"). Variety was assessed with 3 items (eg, "How often do you do a variety of different things on your job?"). Psychological demand was assessed with 4 items tapping the stressors inherent in participants' jobs (eg, "How often is your job hectic?"). Items for each concept were coded and summed such that higher values indicated greater frequency of exposure to the latent concept. The estimated internal consistency of the control, variety, and psychological demand scales was .92, .68 and .66, respectively. Perceived safety commitment was assessed using an existing 7-item scale³⁵ and constructed by coding and summing the items so that higher values indicate greater perceived management commitment to worker safety ($\alpha = .74$). Abusive supervision was measured with a 7-item index assessing the extent to which supervisors/managers use coercive tactics with their employees.³⁶ Items were summed with higher values indi-

cating greater use of coercive tactics by managers/crew chiefs ($\alpha = .75$).

Analysis

The main variable of interest in this analysis is adherence status, which is a discrete categorical variable constructed from the number of days a participant completed a daily diary. Frequencies and percentages were calculated to describe participants in each adherence category. Chi-square tests were used to test for differences among adherence status groups on categorical measures, most of which were demographic characteristics (eg, age-group, marital status, education). One-way analysis of variance and resultant F statistics were used to make comparisons in measures of health at baseline (eg, DASH, CES-D) among the adherence status groups, and logistic regression models were fit to test for differences between adherence status groups on continuous measures of work characteristics (eg, hours of work per week, weeks of construction). Statistical tests were considered significant at $p < .05$.

RESULTS

All participants except for 1 were male, and fully one-third were less than 29 years of age, and over 75% were less than 39 years (Table 1). The majority of participants were married, and their spouses were in the United States. The modal educational attainment was *secundaria*, which is generally comparable to a sixth-to-ninth-grade education in the United States. Over one third of participants had been in the United States for fewer than

Table 3
Baseline Work and Job Characteristics by Adherence Status to the 21-Day IVR Telephone Diary

Characteristics	Sample N (%)	Noncompleters N (%)	Nonadherent N (%)	Adherent N (%)	p-value
Worker Status					.04*
Framer	26 (21.8)	11 (42.3)	8 (30.8)	7 (26.9)	
Roofers	35 (29.4)	7 (20.0)	10 (28.6)	18 (51.4)	
General construction	58 (48.7)	11 (19.0)	28 (48.3)	19 (32.8)	
	M (SD)	M (SD)	M (SD)	M (SD)	
Hours of Work/Week	41.9 (9.0)	43.3 (6.9)	41.3 (6.8)	41.5 (11.8)	.62
Weeks of Construction	37.7 (16.6)	35.6 (15.0)	38.8 (15.5)	37.8 (18.9)	.74
Control	6.2 (3.5)	7.1 (3.7)	6.3 (3.6)	5.5 (3.1)	.15
Variety	8.4 (2.3)	9.2 (2.4)	8.1 (2.1)	8.1 (2.4)	.09
Demand	9.8 (3.1)	9.9 (2.8)	10.0 (3.3)	9.6 (3.2)	.79
Perceived Safety Climate	23.0 (5.3)	24.3 (5.5)	22.4 (5.4)	22.7 (5.1)	.29
Supervisory Abuse	17.2 (4.2)	18.1 (4.0)	17.2 (4.4)	16.6 (4.0)	.31

Note.

* p-value obtained using chi-square test. Remaining p-values obtained from univariate analysis of variance models and corresponding F statistics, wherein the work characteristic was regressed on participant classification.

7 years, and nearly 85% had been in the United States for fewer than 15 years. Over 3 quarters of the sample were from Mexico, whereas the remainder were from another country, primarily Honduras or Guatemala.

One quarter (N=29) of the sample was classified as noncompleters; these participants did not complete one daily diary using the IVR system. Comparable numbers of participants were classified as nonadherent (N=46, 38.7%) and adherent (N=44, 37%). The percentage of noncompleters was greater in the second half of the recruitment and data collection period than in the first half. Otherwise, there were no demographic differences among noncompleters, those who were nonadherent, and those who were adherent.

The average participant received a score of 3.4 (SD = 4.9) on the Disability of the Arm, Shoulder and Hand instrument, and a score of 0.3 (SD = 1.7) on the Rolland-Morris instrument assessing back pain (Table 2). Physical health-related quality of life and mental health-related quality of life were

both below the population norm-based score of 50. Participants had an average depressive symptoms score of 12.9 (SD = 4.8). Importantly, none of health characteristics assessed at the baseline interview differed by adherence to the diary protocol.

There was very little evidence that adherence status to the diary protocol systematically differed by work or job characteristics measured at baseline (Table 3). Framers were more likely to be classified as noncompleters, whereas roofers and general construction laborers were more likely to be classified as adherent and nonadherent, respectively. Otherwise, there were no differences in either structural aspects of construction employment (eg, hours of work/week, weeks in construction) or the way the work is organized (eg, control, psychological demand, safety climate).

Call behavior was systematically different for adherent participants compared to nonadherent participants, as would be expected given the way adherence was defined. Adherent participants completed, on average, 19 (SD = 2.3) telephone calls

Table 4
Agreement With Statements from Debriefing Interview About Using the IVR System (N=85)^a

	Sample (N=85) % (N)	Nonadherent (N=40) % (N)	Adherent (N=45) % (N)	p-value
Appraisals				
The questions I was asked made sense, given the type of work I do.	99% (84)	100% (40)	98% (44)	.53
I liked the confidentiality of answering work-related questions to a computer instead of a person.	72% (61)	75% (30)	69% (31)	.53
The telephone calls were easy.	74% (63)	70% (28)	78% (35)	.41
I would participate in another study like this one that used the telephone calls.	81% (67)	72% (29)	88% (38)	.07
Experienced Barriers				
I had problems because the computer did not understand my response to a question.	64% (54)	60% (24)	67% (30)	.53
I was cut off or had to start the call over for any reason.	38% (32)	38% (15)	38% (17)	.95
I was not able to become comfortable answering questions asked by a computer.	41% (35)	40% (16)	42% (19)	.84
Remembering to call each day was hard for me.	47% (40)	68% (27)	29% (13)	.01
I skipped at least one telephone call because they took too minutes from my cell phone plan.	16% (14)	18% (7)	16% (7)	.81

Note.

a Debriefing interviews were completed with participants who accessed the IVR system at least one time.

during the 21-day period. The average length of the call during workdays was 13.6 minutes (SD = 3.2 minutes), whereas the average call length on nonworkdays was 3.5 minutes (SD = 3.8). Nonadherent participants completed 9 calls, on average, during the 21-day period; and their calls tended to be approximately one minute shorter than calls completed by adherent participants on both workdays and nonworkdays.

During the debriefing interview the vast majority of participants reported that questions asked during the daily diary made sense given the type of work performed and that the telephone calls were perceived as “easy” (Table 4). Over 2 thirds of the sample reported liking the confidentiality of answering work-related questions asked by a computer instead of a person. Nearly 3 quarters of the sample reported that they would participate in another study using the IVR technology; however, trend-level evidence ($p < .10$) suggests that nonadherent participants were less likely to agree with

this statement than were adherent participants. The debriefing interviews also identified areas requiring consideration and attention in future research. Nearly 2 thirds of the sample reported having problems because the IVR system did not understand or recognize the participant’s response to a question, and nearly 40% of participants reported getting cut off and having to restart the interview one or more times. Nearly half of the sample reported difficulty remembering to call each day; this was the only debriefing item that differed significantly between adherent and nonadherent participants. Over 40% of participants reported never becoming comfortable answering questions asked by a computer. Finally, at least one in 7 participants reported purposively skipping at least one telephone call because it took too many minutes from their cell phone plan.

DISCUSSION

Daily diary studies provide valuable insight for

understanding human health behavior, but implementing these types of studies can be challenging in different segments of society.⁶ Literacy and language issues pose significant challenges to collecting high-quality data. IVR technology overcomes many of these challenges, and it has been used to study various behaviors; however, the feasibility of this method among low-literacy Latino immigrants remains understudied. This study addressed this gap, and the results make several contributions to the literature.

IVR technology can be used to study health behaviors in samples of immigrant Latinos from Mexico. Our data indicate that over one third of the sample adhered to the diary protocol using IVR technology, with minimal prompting. Results further indicate that the likelihood of adhering to the protocol does not systematically differ by several demographic, health, and job or occupational characteristics. These results are consistent with previous research supporting the feasibility of IVR technology with non-English-speaking Latino individuals,¹⁶ as well as results suggesting that IVR technology is feasible for collecting data from low-income and other hard-to-reach study populations¹² and for collecting data on sensitive topics (eg, nonheterosexual orientation among specific groups).⁸ Our results extend this research by indicating that IVR technology can be usefully applied in community-based research with non-English-speaking immigrant populations, as well as health behavior research pertaining to occupational safety.

However, future research using IVR technology for data collection with immigrant Latinos should be attentive to 3 distinct results from this study. First, approximately 2 thirds of the study sample did not adhere to the protocol: about 25% never started the diary protocol, and another 38% did not complete at least 85% of the required calls. As researchers advance the use of IVR methods in future studies with immigrant Latinos, attention needs to be given to alternative strategies for promoting adherence to telephone call protocol. Several strategies are advocated to enhance adherence to daily diary-type protocols such as providing complete training on how to use data collection procedures, using reminder devices, and letting participants know that overall adherence will be monitored and taken into consideration when awarding the study incentive.³ These strategies were not emphasized in this study so that we could better evaluate the feasibility of the method, but the results suggest that strategies such as these may be even more salient in this population. Second, our findings reinforce reports of the perils of paper-and-pencil diaries² and one strength of IVR technology; that is, the ability to rigorously evaluate adherence using the time and date stamp of each diary entry.

Our finding that adherence to the diary protocol differed between the first and second half of the recruitment and data collection period is also

useful. We interpret this difference as reflective of the *personalismo* or the strength of the interpersonal relationship of project personnel (ie, the data collector) with study participants. That is, data collectors recruited individuals within their social networks first, and the personal relationship inherent in these social networks enabled better adherence. However, as study-eligible participants within the data collectors' social networks were exhausted, the implicit obligation among individuals with an established social relationship was likely weakened, thereby contributing to poorer adherence. This interpretation is consistent with previous reports of the importance of *personalismo* for recruiting Latinos into research focused on health and related processes³⁷⁻³⁹ and suggests that future studies involving IVR technology with immigrant Latinos will need to complement existing recruitment and retention tactics with those that build in and capitalize on *personalismo*.

Results from the debriefing interviews also provide insight into specific ways to enhance future use of IVR technology for field diary studies with immigrant Latinos. An unacceptably high number of individuals experienced difficulties with the IVR system, both in terms of the number of individuals reporting the system did not understand a response or cut them off during a session. Future applications need to be attentive to subtle features of the VoiceXML coding to enhance users' experience. Two key examples were learned from this project. The first is the importance of hardware and supporting programming that allows differentiation of voice responses from ambient noise. When it is anticipated that many of these calls will take place on a cell phone in the truck on the way home from the job site, the computer needs to differentiate noise from the radio or a rolled-down window from actual voice inputs provided by the respondent to asked questions. Second, this experience highlights the importance of subtle technology and programming logic that affects the sound quality of the VoiceXML application. This is particularly important when speaking Spanish. Once the project was completed, we realized that to fully use foreign characters such as Spanish accents we need to ensure that the VoiceXML be encoded in UTF-8 (Unicode) as opposed to the US ANSI standard encoding character set.

The contributions of this study must be considered in light of its limitations. First, the generalizability of study findings is unknown because the sample was small, regional, and recruited using nonprobability methods. Next, the field of possible predictors of differential adherence to the diary protocol was circumscribed. Although we considered many useful demographic, individual (ie, health), and job-related determinants of adherence, other factors such as documentation or cultural factors were not assessed. Additional research is needed to fully discern the extent to which IVR technology can be used for data collection with specific

subgroups of immigrant Latinos. Next, research is needed to evaluate alternative diary periods to effectively capture workers' health and safety behavior, while also avoiding possible threats to data integrity resulting from diary-induced behavior changes (ie, answering regular questions about behavior may result in different behavior patterns) or habituation to diary surveys over extended periods of time. Finally, this study did not evaluate the overall quality of data obtained from IVR technology. Future research that systematically examines response patterns to individual items and scales across a study period is needed before definitive conclusions can be drawn about using IVR technology for data collection with immigrant Latinos.

Limitations notwithstanding, this study adds to the growing evidence base of the value of IVR as a data collection tool. The evidence suggests that immigrant Latinos can, with minimal prompting, adhere to a 21-day diary protocol and that there are few individual or contextual factors that systematically shape whether or not a participant will adhere. Emphasis on *personalismo* may contribute to enhanced adherence to use of the IVR system. Although additional evaluation research is needed, the overall pattern of results suggests that IVR technology may provide a tool to enable use of daily diary-type study designs for understanding health behaviors among immigrant Latinos.

Human Subjects Statement

All participants provided signed informed consent. The Wake Forest School of Medicine Institutional Review Board (FWA #00001435) approved all sampling, recruitment, and data collection procedures.

Acknowledgment

This research was supported by a grant from the National Institute for Occupational Safety and Health (OH009761-01, subproject #647).

REFERENCES

1. Tennen H, Affleck G, Armeli S, Carney MA. A daily process approach to coping: linking theory, research, and practice. *Am Psychol*. 2000;55(6):626-636.
2. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. *Ann Behav Med*. 2002;24(3):236-243.
3. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. *Annu Rev Psychol*. 2003;54(1):579-616.
4. Burton C, Weller D, Sharpe M. Are electronic diaries useful for symptoms research? A systematic review. *J Psychosom Res*. 2007;62(5):553-561.
5. Grzywacz JG, Almeida DM, Neupert SD, Ettner SL. Socioeconomic status and health: a micro-level analysis of exposure and vulnerability to daily stressors. *J Health Soc Behav*. 2004;45(1):1-16.
6. Abu-Hasaballah K, James A, Asetline J. Lessons and pitfalls of interactive voice response in medical research. *Contemporary Clinical Trials*. 2007;28(5):593-602.
7. Collins RL, Kashdan TB, Gollnisch G. The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Exp Clin Psychopharmacol*. 2003;11(1):73-78.
8. Midanik LT, Greenfield TK. Interactive voice response versus computer-assisted telephone interviewing (CATI) surveys and sensitive questions: the 2005 National Alcohol Survey. *J Stud Alcohol Drugs*. 2008;69(4):580-588.
9. Barta WD, Kiene SM, Tennen H, et al. The idiographic study of inconsistent condom use behavior of persons living with HIV. *AIDS Care*. 2007;19(8):1058-1064.
10. DeRogatis LR, Allgood A, Auerbach P, et al. Validation of a women's sexual interest diagnostic interview--short form (WSID-SF) and a daily log of sexual activities (DLSA) in postmenopausal women with hypoactive sexual desire disorder. *J Sex Med*. 2010;7(2 pt 2):917-927.
11. Beach SR, Schulz R, Degenholtz HB, et al. Using audio computer-assisted self-interviewing and interactive voice response to measure elder mistreatment in older adults: feasibility and effects on prevalence estimates. *Journal of Official Statistics*. 2010;26(3):507-533.
12. Kim HG, Geppert J, Quan T, et al. Screening for postpartum depression among low-income mothers using an interactive voice response system. *Matern Child Health J*. 2011;1-8.
13. Forster AJ, Boyle L, Shojania KG, et al. Identifying patients with post-discharge care problems using an interactive voice response system. *J Gen Intern Med*. 2009;24(4):520-525.
14. Bender BG, Apter A, Bogen DK, et al. Test of an interactive voice response intervention to improve adherence to controller medications in adults with asthma. *J Am Board Family Med*. 2010;23(2):159-165.
15. Naylor MR, Naud S, Keefe FJ, Helzer JE. Therapeutic interactive voice response (TIVR) to reduce analgesic medication use for chronic pain management. *J Pain*. 2010;11(12):1410-1419.
16. Brodey BB, Rosen CS, Brodey IS, et al. Reliability and acceptability of automated telephone surveys among Spanish-and English-speaking mental health services recipients. *Ment Health Serv Res*. 2005;7(3):181-184.
17. ICPWR Data Center. Hispanic Employment in Construction. Available at: <http://www.cpw.com/pdfs/Hispanic-Data-Brief-Nov-09.pdf>. Accessed February 5, 2012.
18. Azaroff LS, Levenstein C, Wegman DH. Occupational injury and illness surveillance: conceptual filters explain underreporting. *Am J Public Health*. 2002;92(9):1421.
19. Menzel NN, Gutierrez AP. Latino worker perceptions of construction risks. *Am J Ind Med*. 2010;53(2):179-187.
20. Quandt SA, Grzywacz JG, Marin A, et al. Illnesses and injuries reported by Latino poultry workers in western North Carolina. *Am J Ind Med*. 2006;49(5):343-351.
21. Kochhar R, Suro R, Tafoya S. The New Latino South: The Context and Consequences of Rapid Population Growth," Pew Hispanic Center, 2005. Available at: <http://www.pewhispanic.org/files/reports/50.pdf>. Accessed April 5, 2012.
22. Behling O, Law KS. Translating questionnaires and other research instruments: Problems and solutions. Sage Publications, Inc, 2000.
23. Rosales RS, Delgado EB, Lastra-Bosch ID. Evaluation of the Spanish version of the DASH and carpal tunnel syndrome health-related quality-of-life instruments: cross-cultural adaptation process and reliability. *J Hand Surg Am*. 2002;27(2):334-343.
24. Kovacs FM, Llobera J, Gil del Real MT, et al. Validation of the Spanish version of the Roland-Morris questionnaire. *Spine*. 2002;27(5):538-542.
25. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry low back pain disability questionnaire and the Quebec back pain disability scale. *Phys Ther*. 2001;81(2):776-788.
26. Ware JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 health survey: standard & acute forms. Lincoln, RI: QualityMetric Inc., 2000.
27. Radloff LS. The CES-D scale. *Appl Psychol Meas*.

- 1977;1(3):385-401.
28. Hovey JD. Psychosocial predictors of depression among Central American immigrants. *Psychol Rep.* 2000;86(3 pt 2):1237-1240.
 29. Golding JM, Aneshensel CS, Hough RL. Responses to Depression Scale items among Mexican-Americans and non-Hispanic whites. *J Clin Psychol.* 1991;47(1):61-75.
 30. Grzywacz JG, Hovey JD, Seligman LD, et al. Evaluating short-form versions of the CES-D for measuring depressive symptoms among immigrants from Mexico. *Hisp J Behav Sci.* 2006;28(3):404-424.
 31. Hovey JD, Magaña CG. Suicide risk factors among Mexican migrant farmworker women in the Midwest United States. *Arch Suicide Res.* 2003;7(2):107-121.
 32. Karasek R, Theorell T. *Healthy work: Stress, Productivity, and the Reconstruction of Working Life.* Basic Books, 1994.
 33. Grzywacz JG, Arcury TA, Marin A, et al. The organization of work: Implications for injury and illness among immigrant Latino poultry-processing workers. *Arch Environ Occup Health.* 2007;62(1):19-26.
 34. Grzywacz JG, Quandt SA, Arcury TA. Immigrant farmworkers' health-related quality of life: an application of the job demands-control model. *J Agric Saf Health.* 2008;14(1):79-92.
 35. Gillen M, Baltz D, Gassel M, et al. Perceived safety climate, job demands, and coworker support among union and nonunion injured construction workers. *J Safety Res.* 2002;33(1):33-51.
 36. Tepper BJ. Consequences of abusive supervision. *Acad Manage J.* 2000;43(2):178-190.
 37. Haack LM, Gerdes AC, Cruz B, Schneider BW. Culturally-modified recruitment strategies for Latino families in clinical child research: a critical first step. *J Child Fam Stud.* 2011;1-7.
 38. Kao B, Lobato D, Grullon E, et al. Recruiting Latino and non-Latino families in pediatric research: considerations from a study on childhood disability. *J Pediatr Psychol.* 2011;36(10):1093-1101.
 39. Larkey LK, Gonzalez JA, Mar LE, Glantz N. Latina recruitment for cancer prevention education via Community Based Participatory Research strategies. *Contemp Clin Trials.* 2009;30(1):47-54.