Customizing Survey Instruments and Data Collection to Reach Hispanic/Latino Adults in Border Communities in Texas

Michelle O’Hegarty, PhD, Linda L. Pederson, PhD, Stacy L. Thorne, PhD, MPH, Ralph S. Caraballo, PhD, MPH, Brian Evans, BA, Leslie Athey, MS, MS, and Joseph McMichael, BS

The size of the Hispanic population in the United States has changed dramatically over the past 35 years. In 1970, 9.6 million Hispanic adults and children resided in the United States (4.7% of the total population); by July 2006, the number had more than quadrupled, reaching 44.3 million Hispanic adults and children, 14.8% of the total adult population. The growth rate of the Hispanic population was more than 3 times the growth rate of the entire US population between 2000 and 2006. In 2004, the Census Bureau projected that the number of Hispanics living in the United States will reach 102.6 million by the year 2050 and will likely compose about 24% of the population.

There is reason to believe that this population may not be readily accessible for data collection. Documentation of some of the potential barriers to participation can be found in ethnographic research carried out during the 2000 Census. Ethnographic studies in colonias (residential areas along the Texas–Mexico border that may lack some basic necessities), one of which was in Texas, documented barriers to census enumeration: irregular housing, little or no knowledge of English, limited formal education, concerns regarding confidentiality, and complex and fluid households. Translations, enumeration strategies, advertisements, and reassurance about confidentiality were employed, and resulting evidence reveals success in addressing the concerns. The importance of community collaboration for improving participation rates has been documented.

One health-related behavior of interest for all populations is tobacco use. Hispanic and Latino populations in the United States face unique challenges that may put them at higher risk for tobacco use and for exposure to secondhand smoke for several reasons: (1) racial/ethnic communities continue to be targeted by tobacco industry sponsorship efforts, including Mexican rodeos, Cinco de Mayo (May 5th) festivities, and activities associated with Hispanic heritage month; (2) more Hispanic and Latino workers may be exposed to secondhand smoke in the workplace than are other groups, despite the fact that there is a high percentage of smoke-free homes and strong support for smoke-free policies in workplaces among Hispanics and Latinos; (3) preliminary findings suggest that as Hispanic and Latino women become more acculturated into mainstream US society their smoking rates may increase, a pattern that does not seem to be occurring among men, where the rate is already elevated; and (4) overall prevalence rates of current smoking (smoking every day or some days) among Hispanic adults remains low at 13.3% and among Hispanic high-school students (smoking at least once in the past 30 days) is 16.7%. However, there may be marked differences in prevalence rates among subgroups of Hispanics and Latinos. For example, a recent analysis of 2002–2005 data from the National Survey on Drug Use and Health revealed that for Mexican respondents, the prevalence of past-30-day cigarette smoking was 23.8%, whereas for those self-identified as Puerto Rican the prevalence was 31.5%.

Some research has revealed that the attitudes, behaviors, knowledge, and experience of Hispanic and Latino persons residing in the United States regarding tobacco use may differ from those of persons in non-Hispanic groups, which may warrant customized approaches to smoking prevention and cessation programs. Although Hispanic young adults acknowledge the adverse health effects associated with tobacco use, smoking among Hispanic youths continues to be a social activity. In Hispanic cultures, smoking among females is actively discouraged. Young adults also appear to be unwilling to access resources for cessation partly because family and family relationships are considered to be an integral part of smoking prevention and cessation. Targeted, sustained...
Interventions in the Spanish language may also be limited, in part because of the perception that tobacco use is not a problem among Hispanic and Latino populations.

A strategy that may help identify and address tobacco-related issues among Hispanic populations in the United States is the collection of specific data measuring knowledge, attitudes, beliefs, and behaviors that is obtained in culturally sensitive and relevant ways. To this end, M.O., L.L.P., S.L.T., and R.S.C., who were affiliated with the Centers for Disease Control and Prevention’s (CDC’s) Office on Smoking and Health, examined available survey methods, tobacco-related instruments, and their utility for obtaining information from Hispanic populations. An additional objective was to achieve acceptable levels of participation and response rates from a population that does not always respond to standard telephone or mail survey methods.3 There were 2 phases to the project; the first focused on designing a Hispanic/Latino Adult Tobacco Survey (Hispanic/Latino ATS) to collect information from this population. The second phase of the project involved developing data collection methods and sampling techniques that would reach the target population (Hispanic and Latino adults living in the United States).

METHODS

In April 2002, a meeting on Effective Tobacco Control in Hispanic/Latino communities was held by the CDC and was attended by state tobacco control partners and CDC staff. (State participants were selected on the basis of their work with specific Hispanic and Latino populations in the areas of surveillance and evaluation, program, policy, communication, and community advocacy.) The discussion covered a broad range of topics, including suggestions of how to categorize different population subgroups, protective factors (against tobacco use) that could be operating in Hispanic and Latino populations, the cultural context for cigarette use and use of noncigarette tobacco products such as smokeless tobacco, as well as the role of acculturation in the health behaviors of this population. There was also discussion of the methods used for data collection and sampling. The conclusions from the meeting provided the basis for revisions made to the State Adult Tobacco Survey (State ATS) to create the Hispanic/Latino ATS.

Hispanic/Latino Adult Tobacco Survey

The modifications to the State ATS were made keeping in mind the monitoring and evaluation needs of individual states and comparability across states. The modifications were designed to incorporate measures and language to capture attitudes, behaviors, knowledge, and experience of these specific subpopulations and to be sensitive to the varied and unique cultural and social experiences of Hispanic and Latino persons. Since July 2009, 25 states have completed at least 1 State ATS supported by the Office on Smoking and Health at CDC (S. Thorne; Health Scientist, Office on Smoking and Health, CDC; written communication; August 2009). The objective was to prepare a battery of questions that measured Hispanic and Latino smokers’ and nonsmokers’ (1) smoking behavior and experiences and (2) knowledge, attitudes, beliefs, and behaviors concerning smoking.

Five of the 10 participants in the surveillance and evaluation group of the 2002 expert advisory panel were selected to participate in the revision of the instrument on the basis of their expertise in tobacco control and their research experience with Hispanic and Latino subpopulations (e.g., Mexican American, Puerto Rican, Cuban, Central American, and South American). Following the revisions, 2 teams of professional translators translated and cognitively tested the instrument in English and Spanish.15–18 They reconciled discrepancies and reached agreement on a translation that best achieved equivalence of measurement between English and Spanish and between different varieties and dialects of Spanish. In 2004 and 2005, the questionnaire was cognitively tested in Chicago, Illinois; Los Angeles, California; Miami, Florida; New York, New York; San Antonio, Texas; El Paso, Texas; and Washington, DC.

Sixty-eight interviews were conducted with Hispanic and Latino respondents (16 in English and 52 in Spanish).19 Several subpopulations were included in the cognitive testing of the questionnaire (e.g., respondents from Mexico, Puerto Rico, Cuba, Guatemala, El Salvador, Columbia, Dominican Republic, Peru, Ecuador, and Honduras).

Sample Design and Data Collection

From July 2007 to April 2008, the Hispanic/Latino ATS was administered in the southeast colonias in El Paso. (Research Triangle Institute International was contracted by the CDC to complete the field portion of these interviews. Research Triangle Institute International contracted with Texas A&M University to complete the data collection for the study.) According to the Texas Secretary of State Web site, a colonia is “a residential area along the Texas–Mexico border that may lack some of the most basic living necessities, such as potable water and sewer systems, electricity, paved roads, and safe and sanitary housing.”20–22 Over the past 15 years, Texas has adopted several laws to address infrastructure problems associated with colonias, thereby ensuring water and sewer services and limiting utility connections in substandard developments.22 The sample frame of colonias was grouped into 2 categories (green and not-green) depending on the level of public services available. “Green” colonias received the highest level of services, and “not-green” colonias were those with lower or unknown levels of services.

Because of the experiences in the 2000 US Census and their past experience in working with residents of the colonias, individuals from the Texas A&M University Colonias Program administered the survey. The interviewers met the program’s criteria for promotoras (outreach workers who assist in the development of community programs in the colonias). There were 9 female interviewers and 1 male interviewer, who ranged in age from early 20s to late 40s. An intense 3-day training program was conducted in Spanish and covered the following topics: background of the study, interviewer responsibilities, interviewing skills and techniques, locating and sampling participants, and questionnaire administration. Role playing was used to ensure that the interviewers were proficient in the required skills.

Mailing lists could not be used as the sampling frame because post office boxes were widely used in the area. Post office boxes are not suitable for in-person surveys because they are not locatable on the ground. Instead, the project team used the GeoFrame (RTI International, Research Triangle Park, NC) field enumeration process; this involved the use of...
digital photography and geospatial technology to produce a low-cost household enumeration of all of the colonias. With this method, the project enumeration team photographed each dwelling unit with a digital camera, and identified global positioning system coordinates. As a result, the project team created the sampling frame from a complete enumeration of the 8990 apparent dwelling units in the defined area, and dwelling units were randomly selected from this enumeration. The project team provided the interviewers with the geospatial information and a photograph of the selected unit.

For the purposes of this survey, the project team defined a dwelling unit as a single structure containing 1 or more rooms or a trailer, converted bus, or other domicile where people could sleep. It was not unusual in the colonias to find multiple dwelling units on a single property that housed various branches of a single family or to find more than 1 household occupying a single dwelling unit. A screening interview was used to determine the number of Hispanic adults (aged 18 years or older) who lived in the household and to select the adult with the latest birthday as the respondent. If no one answered the door, a note in English and Spanish describing the study, the purpose of the visit, and a telephone number for scheduling an appointment was left.

RESULTS

The survey is made up of 6 core sections that include questions about cigarette smoking, cessation, secondhand smoke, risk perception, and social influences, demographic items, and supplemental questions to assess specific program needs. Optional modules can be added to the core to obtain additional information on use of noncigarette tobacco products, cessation, secondhand smoke exposure, health and social influences, policy issues, parental involvement, media exposure, and additional demographic items. The revised questionnaire included 36 unchanged items from the State ATS, 7 items modified with word choice changes or ordering, and 17 new items that focused on possible culturally specific methods used for quitting (e.g., medicine man, spiritist, or religious leader), secondhand smoke issues, and attitudes and knowledge about tobacco use that might be unique for Hispanic and Latino groups. (The resulting instruments in English and Spanish are available at http://www.cdc.gov/tobacco/data_statistics/surveys/hispanic_latino_ats_guide.)

The Sample

The project team selected the sample in 2 stages (see Table 1); the first stage consisted of a random sample of 2100 apparent dwelling units in each of the 2 strata (green or not-green) of the colonias. The sample of units was distributed as evenly as possible between the 2 strata of green or not-green colonias. The project team based eligibility on the selected unit being a dwelling unit, and self-report of Hispanic or Latino residents, with at least 1 resident who was aged 18 years or older. During data collection, the interviewers determined that 203 units of the selected sample were ineligible because they were not dwelling units, thus leaving 1897 units. Interviewers were able to screen household members in 1713 of these. In the second stage, the interviewers selected 1 adult from each of the eligible dwelling units with the “last birthday” method. There was no cap placed on the number of visits to a specific household. Instead, in the daily review of eligible households, the interviewers and project team identified the households that were difficult to contact in-person and attempts were made at different times of day and different days of the week. After 15 in-person visits, the data collection manager attempted to make phone contact for 10 households where a phone number was obtained.

Data Collection

The survey was conducted from July 2007 through April 2008; Office of Management and Budget clearance was obtained by CDC. Both Research Triangle Institute International and Texas A&M University’s institutional review boards approved the study. The screening interview took an average of 5 minutes to complete; the main interview took an average of 21 minutes (ranging from 8 to 60 minutes). In 29 cases, when face-to-face interviews could not be scheduled, the interview was conducted by phone. Participants who completed the main interview received a $15 Wal-Mart gift card for their time and participation. The project team maintained quality control in the field by (1) having the study coordinator review all questionnaires so that immediate feedback could be given to the interviewers; (2) having the study coordinator visit 5% to 9% of each team’s cases to verify that the interview had occurred and to collect information on the team’s professionalism (there were no instances of false reporting by the teams); and (3) having the data collection manager periodically visit to observe interviews and monitor procedures.

Response Rates

Table 2 displays the final status of the 2100 cases. The calculation of the response rate with

<table>
<thead>
<tr>
<th>Table 1—Distribution of the Sample: Colonias in El Paso, Texas, 2007–2008</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage of Sampling</strong></td>
</tr>
<tr>
<td><strong>First stage: enumerated dwelling units (DUs)</strong></td>
</tr>
<tr>
<td>Sampling frame</td>
</tr>
<tr>
<td>Stratification</td>
</tr>
<tr>
<td>Type of selection</td>
</tr>
<tr>
<td>Sample sizes</td>
</tr>
<tr>
<td><strong>Second stage: eligible persons</strong></td>
</tr>
<tr>
<td>Sampling frame</td>
</tr>
<tr>
<td>Stratification</td>
</tr>
<tr>
<td>Type of selection</td>
</tr>
<tr>
<td>Sample sizes</td>
</tr>
</tbody>
</table>

*We confirmed 203 dwelling units were not eligible because of vacancy, being a business, or being demolished. Of the remaining 1897 cases, we were unable to contact a household member for 184 of them. The interviewers visited each of these 184 dwelling units. Because they could not confirm they were not dwelling units, we assumed they were eligible dwelling units.
the method of the American Association for Public Opinion Research can be found in Table 3. American Association for Public Opinion Research response rate 228 provides a conservative calculation, because it assumes all of the unknown cases were eligible households. The eligibility rate was 90.2%, the conservative combined screener and interview response rate was 80.0%, and the hit rate (the percentage of all cases released that resulted in a completed interview) was 70.7%. Costs, which included enumeration of dwelling units, random selection of the units and the respondent, completion of the interview, and entry of the responses, were estimated to be less than $300 per completed interview. Cost information for instrument development was not available.

**DISCUSSION**

The results support the conclusion that culturally sensitive modifications to survey procedures used to locate and contact specific population groups can result in response rates that far exceed those common in survey work today.26,29,30 The 80% combined screener and interview response rate in the current survey is not typical of response rates in other survey work. For example, the 2007 Behavioral Risk Factor Surveillance System, which used telephone and not personal interviews, reported response rates from 26.90% to 65.36%.31 Direct comparisons of the response rates and information obtained with these methods need to be made with caution because of the many differences in methodologies.

Developing instruments for specific population subgroups requires consideration of culture and language, cognitive demands, and potential response errors.32 Collecting information from specific subpopulations requires community knowledge and specialized training. The interviewers were trained to collect the information in a manner that was culturally sensitive and appropriate. Some of these interview methods focused on, but were not limited to, conducting the survey in Spanish, cognizance of demeanor during the interview, tone, and probing techniques. There was a considerable amount of monitoring of the methods and interaction with interviewers by the project staff from the Texas A&M Colonias Program. This approach is ideal for collecting information from specific subpopulation groups. What has been demonstrated is that a strategic, targeted, carefully designed survey can result in high response rates. Whether this rate can be achieved in other locations is not known until the findings from the present study are replicated.

The solutions to the numerous data collection challenges encountered on the Hispanic/Latino ATS may be applicable to surveys of other population subgroups. GeoFrame was developed to assist in the creation of an area probability sample in which there was little or no landline telephone coverage and a lack of postal listings. A question may arise as to whether it is a cost-effective and timely solution to create a sample similar to the one used in this study. For many areas in the United States that do not have postal listings or landline telephone coverage, such an approach may be considered. Responses to this question may vary depending on population size and density. The utility of the approach is not limited to the United States; it may be useful in international surveys where there are not already existing sampling frames or may extend to surveys conducted in the aftermath of war or natural disasters, when up-to-date residential listings are unavailable.

To overcome barriers to clear and concise communication, a great deal of preliminary work was done in the development and modification of the questionnaire. Numerous field tests of the survey were conducted, along with input from stakeholders and target population, before the actual implementation of the survey. Although this may not be possible for all surveys, at a minimum close attention should be paid to the development and translation of the questionnaire, taking into account cultural considerations as well as verbal or written comprehension by the target population.

When we began preparations for this survey, we assumed that telephone interviews were not appropriate for the target population. We designed an interview scenario in which respondents would feel comfortable by hiring native Spanish-speaking interviewers who were familiar with the interview area. In addition, training methods were used to allow interviewers to gain experience in interviewing procedures that were used in collecting survey data. Finally, measures were taken to assist interviewers in overcoming potential objections by respondents. Involvement of the Texas A&M Colonias Program proved vital to the success of the study, as they had insight and understanding of the colonias and their residents that proved to be essential in ensuring cooperation. The use of uniforms with logos from the Texas A&M and Texas A&M Colonias Program and other Colonias Program materials may have helped to ease concerns for respondents about strangers in their community.

**Strengths and Limitations**

There are several important strengths in this study; first is the use of a survey that was designed to be culturally appropriate for Hispanic and Latino persons; second, the use of promotoras who were Spanish-speaking and familiar with the interview area; and, third, the use of GeoFrame to create a probability sample in which traditional methods could not be used because of inconsistent telephone coverage and postal listings. The approach to sampling can be applied in other geographic areas where sampling frames are not readily available such as in geographic regions that have been devastated by hurricanes or in isolated rural regions in countries after other catastrophic events such as war.

There were also some limitations: the use of personal interviews may have resulted in interviewer bias if respondents were uncomfortable communicating sensitive personal

---

**TABLE 2—Disposition of the Sample: Colonias in El Paso, Texas, 2007–2008**

<table>
<thead>
<tr>
<th>Description</th>
<th>No. of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final screener codes</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2100</td>
</tr>
<tr>
<td>Completed screener</td>
<td>1713</td>
</tr>
<tr>
<td>Unknown eligibility—no contacts, unable to locate</td>
<td>184</td>
</tr>
<tr>
<td>Not eligible—vacant, business, demolished units</td>
<td>203</td>
</tr>
<tr>
<td>Final interview codes</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1713</td>
</tr>
<tr>
<td>Completed interviews</td>
<td>1485</td>
</tr>
<tr>
<td>Eligible, noninterview—refusal, unable to contact</td>
<td>187</td>
</tr>
<tr>
<td>Not eligible—Hispanic, younger than 18 years, incapable</td>
<td>41</td>
</tr>
</tbody>
</table>
In summary, this work was carried out to tailor a survey to gather population-specific information and increase survey participation from hard-to-reach populations. It is not possible to determine which components of the survey modifications and methods used to collect the data had the greatest impact on response rates in this study, nor whether these efforts could have worked as well in isolation from each other. However, the comprehensive approach used here provided response rates that exceeded those occurring in other surveys of this population and may be useful in other hard-to-reach communities.

About the Authors
At the time of the study, Michelle O’Hegarty, Linda L. Pederson, Stacy L. Thorne, and Ralph S. Caraballo were with the Office and Smoking and Health, Centers for Disease Control and Prevention (CDC), Atlanta, GA. Brian Evans, Leslie Athey, and Joseph McMichael were with Research Triangle International, Research Triangle Park, NC.

Correspondence should be sent to Michelle O’Hegarty, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mailstop K-50, Atlanta, GA 30341-3717 (e-mail: mohegarty@cdc.gov). Reprints can be ordered at http://www.ajph.org by clicking the “Reprints/Eprints” link.

This article was accepted August 25, 2009.

Note. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC. Use of trade names is for identification only and does not imply endorsement by the US Department of Health and Human Services.

Contributors
M. O’Hegarty was the CDC’s principal investigator for the Hispanic/Latino Adult Tobacco Survey (Hispanic/Latino ATS). She was responsible for leading the preparation of this article and contributed to the writing of the article. L. L. Pederson was responsible for writing the initial and final drafts of the article. S. Thorne contributed to the writing of the article and facilitated the edits of the article requested through internal review at CDC. R. S. Caraballo was the CDC senior advisor on the Hispanic/Latino ATS project and contributed to the writing of the article. B. Evans participated in the design of the methodology used in the survey, the administration of the survey in El Paso, Texas, and the writing of the article. J. McMichael contributed to the sample design and the analysis. L. Athey was the project staff manager in El Paso; she led the survey administration team and contributed to preparation of the article.

Acknowledgments
The authors wish to thank the assistance of David Holiday and Victoria Albright of Research Triangle Institute International for their assistance with the study and sample design, Rebecca Zima of the Texas Department of State Health Services for her guidance on selection of the colonias, Pemá Garcia of the Western Region for the Colonias Program for her advice on the study design and data collection, and the Colonias Program staff of the Texas A&M University Center for Housing and Urban Development for conducting the interviews. We also wish to thank the study participants for their time and contributions, the reviewers for their comments, and Akiko Wilson of CDC for her graphics.
assistance. We also want to acknowledge Alisl Schoua-Glusberg of Research Support Services, Elva Yanez, an independent consultant, Isabel Scarnici of the University of Alabama at Birmingham, Jesse Nodora of the Family and Community Medical College of Medicine at Arizona Cancer Center; Lourdes E. Soto de Laurido, associate dean and director for the Research Institute for Global Health Promotion and Health Education, School of Health Professions; and Maribel Quiala, an independent consultant, for their assistance with the Hispanic/Latino ATS revisions and development.

Human Participant Protection
This study was approved by the institutional review board at Research Triangle Institute International and Texas A&M University.

References